

Seq-15 2 w/ Inters

OM protein - protein search, using sw model
Run on: December 17, 2004, 21:53:48 ; Search time 466 Seconds
(without alignments)
1581.157 Million cell updates/sec

Title: US-10-606-607-2
Perfect score: 3479
Sequence: 1 MSEKNFYITPIYPSGKLH.....DGKLTLLTVDPAVNGSVIG 665

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6730630 seqs, 1107998698 residues

Total number of hits satisfying chosen parameters: 6730630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata1/paa/PCTUS COMB.pcp.*
- 2: /cgn2_6/ptodata1/paa/US06 COMB.pcp.*
- 3: /cgn2_6/ptodata1/paa/US07 COMB.pcp.*
- 4: /cgn2_6/ptodata1/paa/US08 COMB.pcp.*
- 5: /cgn2_6/ptodata1/paa/US081 COMB.pcp.*
- 6: /cgn2_6/ptodata1/paa/US082 COMB.pcp.*
- 7: /cgn2_6/ptodata1/paa/US083 COMB.pcp.*
- 8: /cgn2_6/ptodata1/paa/US084 COMB.pcp.*
- 9: /cgn2_6/ptodata1/paa/US085 COMB.pcp.*
- 10: /cgn2_6/ptodata1/paa/US086 COMB.pcp.*
- 11: /cgn2_6/ptodata1/paa/US087 COMB.pcp.*
- 12: /cgn2_6/ptodata1/paa/US088 COMB.pcp.*
- 13: /cgn2_6/ptodata1/paa/US089 COMB.pcp.*
- 14: /cgn2_6/ptodata1/paa/US090 COMB.pcp.*
- 15: /cgn2_6/ptodata1/paa/US091 COMB.pcp.*
- 16: /cgn2_6/ptodata1/paa/US092 COMB.pcp.*
- 17: /cgn2_6/ptodata1/paa/US093 COMB.pcp.*
- 18: /cgn2_6/ptodata1/paa/US094 COMB.pcp.*
- 19: /cgn2_6/ptodata1/paa/US095 COMB.pcp.*
- 20: /cgn2_6/ptodata1/paa/US096 COMB.pcp.*
- 21: /cgn2_6/ptodata1/paa/US097A COMB.pcp.*
- 22: /cgn2_6/ptodata1/paa/US097B COMB.pcp.*
- 23: /cgn2_6/ptodata1/paa/US098 COMB.pcp.*
- 24: /cgn2_6/ptodata1/paa/US099A COMB.pcp.*
- 25: /cgn2_6/ptodata1/paa/US099B COMB.pcp.*
- 26: /cgn2_6/ptodata1/paa/US100 COMB.pcp.*
- 27: /cgn2_6/ptodata1/paa/US101 COMB.pcp.*
- 28: /cgn2_6/ptodata1/paa/US102 COMB.pcp.*
- 29: /cgn2_6/ptodata1/paa/US103 COMB.pcp.*
- 30: /cgn2_6/ptodata1/paa/US104 COMB.pcp.*
- 31: /cgn2_6/ptodata1/paa/US105 COMB.pcp.*
- 32: /cgn2_6/ptodata1/paa/US106 COMB.pcp.*
- 33: /cgn2_6/ptodata1/paa/US107 COMB.pcp.*
- 34: /cgn2_6/ptodata1/paa/US108 COMB.pcp.*
- 35: /cgn2_6/ptodata1/paa/US109 COMB.pcp.*
- 36: /cgn2_6/ptodata1/paa/US60 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3479	100.0	665	1	PCT-US97-06555-2
2	3479	100.0	665	1	Sequence 2, Appli
3	3479	100.0	665	18	Sequence 2, Appli
4	3479	100.0	665	32	Sequence 2, Appli
5	3446	99.1	665	30	Sequence 1498, Ap
6	3446	99.1	679	1	Sequence 13418, A
7	3446	99.1	679	23	Sequence 13416, A
8	3446	99.1	679	26	Sequence 13418, A
9	3446	99.1	679	28	Sequence 73893, A
10	3429	98.6	679	32	Sequence 2751, A
11	2883.5	82.9	666	28	Sequence 74415, A
12	2883.5	82.9	666	30	Sequence 2228, Ap
13	2883	82.9	669	28	Sequence 72423, A
14	2732.5	78.5	665	30	Sequence 10044, A
15	2732.5	78.5	682	30	Sequence 2226, Ap
16	2295	66.0	669	28	Sequence 57809, Ap
17	2291.5	65.9	674	30	Sequence 6201, Ap
18	2291.5	65.9	674	30	Sequence 6201, Ap
19	2288	65.8	677	1	Sequence 10663, A
20	2288	65.8	677	23	Sequence 10663, A
21	2288	65.8	677	26	Sequence 10663, A
22	2288	65.8	677	28	Sequence 42475, A
23	2284.5	65.7	666	1	Sequence 4933, Ap
24	2284.5	65.7	666	23	Sequence 4933, Ap
25	2284.5	65.7	666	26	Sequence 4933, Ap
26	2284.5	65.7	666	36	Sequence 4701, Ap
27	2154	61.9	680	1	Sequence 2080, Ap
28	2154	61.9	680	31	Sequence 2080, Ap
29	2113	60.7	664	28	Sequence 60385, A
30	2040	58.6	652	1	Sequence 2078, Ap
31	2040	58.6	652	31	Sequence 2078, Ap
32	1963.5	56.4	664	22	Sequence 97285, A
33	1949	56.0	659	28	Sequence 71524, A
34	1935.5	55.6	669	28	Sequence 46123, A
35	1934	55.6	657	1	Sequence 12135, A
36	1934	55.6	657	23	Sequence 12135, A
37	1934	55.6	657	26	Sequence 12135, A
38	1931	55.5	657	11	Sequence 12135, A
39	1931	55.5	657	13	Sequence 2, Appli
40	1931	55.5	657	17	Sequence 2, Appli
41	1931	55.5	657	28	Sequence 43933, A
42	1931	55.5	657	34	Sequence 559, App
43	1931	55.5	657	36	Sequence 559, App
44	1930	55.5	657	1	Sequence 13013, A
45	1930	55.5	657	23	Sequence 13013, A

ALIGNMENTS

RESULT 1
PCT-US97-06555-2
; Sequence 2, Application PC/TUS9706555
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: Novel Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; City: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US97/06555
 FILING DATE: 18-APR-1997
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 9607999.1
 FILING DATE: 16-APR-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Gimmi, Edward R
 REGISTRATION NUMBER: 38,891
 REFERENCE/DOCKET NUMBER: P31456
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-4478
 TELEFAX: 610-270-5090
 TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 665 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US97-06555-2

Query Match 100.0%; Score 3479; DB 1; Length 665;
 Best Local Similarity 100.0%; Pred. No. 1.8e-304;
 Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSEKNFYITTPYPSGKLHIGSAITTIACDVLARYKRLMGVDVFLTGLDGHGQKIQQK	60
Db	1	MSEKNFYITTPYPSGKLHIGSAITTIACDVLARYKRLMGVDVFLTGLDGHGQKIQQK	60
Qy	61	AEAGITPQAYVDGMVGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL	120
Db	61	AEAGITPQAYVDGMVGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL	120
Qy	121	GEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQDR	180
Db	121	GEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQDR	180
Qy	181	LVEFFKAHPEFIPDGRLEMLNFTPEGLDLAVSRFTTWGVPVSPKPVVWIDA	240
Db	181	LVEFFKAHPEFIPDGRLEMLNFTPEGLDLAVSRFTTWGVPVSPKPVVWIDA	240
Qy	241	LLNYATGALGAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVLPDLIAH	300
Db	241	LLNYATGALGAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVLPDLIAH	300
Qy	301	GWFMKDGKMSKGNVVPYEMLVYGLDPLRYLMLNLPVSGDGTFTPDYVGRINYE	360
Db	301	GWFMKDGKMSKGNVVPYEMLVYGLDPLRYLMLNLPVSGDGTFTPDYVGRINYE	360
Qy	361	LANDLGNLNRITVSMINKYFDGQIPAYVEGVTEFFDHLAEVAKSIADPHTHMEAVDYPR	420
Db	361	LANDLGNLNRITVSMINKYFDGQIPAYVEGVTEFFDHLAEVAKSIADPHTHMEAVDYPR	420
Qy	421	ALBAVWTLISRTNKYIDETAPWVLDKDEALROLQASVNSHWQASIRVVAHLIEPFMETS	480
Db	421	ALBAVWTLISRTNKYIDETAPWVLDKDEALROLQASVNSHWQASIRVVAHLIEPFMETS	480
Qy	481	RAVLTOIGLEEVSSLENLSLADPADVTVAKGTPIFPRLNMEELIAYIKEQWGNKPAV	540
Db	481	RAVLTOIGLEEVSSLENLSLADPADVTVAKGTPIFPRLNMEELIAYIKEQWGNKPAV	540
Qy	541	EKEWNPDEVELKLNKDEIKPEDPKVEIRVAEYKEVSKVEGSKLQFLDAGDGEDRQI	600
Db	541	EKEWNPDEVELKLNKDEIKPEDPKVEIRVAEYKEVSKVEGSKLQFLDAGDGEDRQI	600
Qy	601	LSGIATYPNEQELVGGKQVIVANLPRKMMKXYUOGMILSAEHGDKLTLLTVDPAVFN	660
Db	601	LSGIATYPNEQELVGGKQVIVANLPRKMMKXYUOGMILSAEHGDKLTLLTVDPAVFN	660
Qy	661	GSVIG 665	

Db 661 GSVIG 665
 RESULT 2
 PCT-US97-06555A-2
 ; Sequence 2: Application PC/TUS9706555A
 ; GENERAL INFORMATION:
 ; APPLICANT: Lawlor, Elizabeth
 ; TITLE OF INVENTION: Novel Compounds
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Smithline Beecham Corporation
 ; STREET: 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19046
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US97/06555A
 ; FILING DATE: 18-APR-1997
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 9607999.1
 ; FILING DATE: 16-APR-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gimmi, Edward R
 ; REGISTRATION NUMBER: 38,891
 ; REFERENCE/DOCKET NUMBER: P31456
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-270-4478
 ; TELEFAX: 610-270-5090
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 665 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US97-06555A-2

Query Match 100.0%; Score 3479; DB 1; Length 665;
 Best Local Similarity 100.0%; Pred. No. 1.8e-304;
 Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSEKNFYITTPYPSGKLHIGSAITTIACDVLARYKRLMGVDVFLTGLDGHGQKIQQK	60
Db	1	MSEKNFYITTPYPSGKLHIGSAITTIACDVLARYKRLMGVDVFLTGLDGHGQKIQQK	60
Qy	61	AEAGITPQAYVDGMVGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL	120
Db	61	AEAGITPQAYVDGMVGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL	120
Qy	121	GEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQDR	180
Db	121	GEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQDR	180
Qy	181	LVEFFKAHPEFIPDGRLEMLNFTPEGLDLAVSRFTTWGVPVSPKPVVWIDA	240
Db	181	LVEFFKAHPEFIPDGRLEMLNFTPEGLDLAVSRFTTWGVPVSPKPVVWIDA	240
Qy	241	LLNYATGALGAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVLPDLIAH	300
Db	241	LLNYATGALGAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVLPDLIAH	300
Qy	301	GWFMKDGKMSKGNVVPYEMLVYGLDPLRYLMLNLPVSGDGTFTPDYVGRINYE	360
Db	301	GWFMKDGKMSKGNVVPYEMLVYGLDPLRYLMLNLPVSGDGTFTPDYVGRINYE	360

QY 361 LANDLGNLNRVTSMINKYFDGOIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDYP 420
DB 361 LANDLGNLNRVTSMINKYFDGOIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDYP 420
QY 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDQASVMSHWQASIRVVAHLIEPFMMETS 480
DB 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDQASVMSHWQASIRVVAHLIEPFMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPFPRLNMBEETAYIKEQMEGNKPAV 540
DB 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPFPRLNMBEETAYIKEQMEGNKPAV 540
QY 541 EKWNPDVELKLNKDEIKFEDPKVEIRVAEVEKSVKVEGSDKLIQFLDAGDGEDRQI 600
DB 541 EKWNPDVELKLNKDEIKFEDPKVEIRVAEVEKSVKVEGSDKLIQFLDAGDGEDRQI 600
QY 601 LSGIAKYYPNEQELVKKVQIVANLKPRKMMKVVQSGMILSAEHDGKLTLLTVDPAV 660
DB 601 LSGIAKYYPNEQELVKKVQIVANLKPRKMMKVVQSGMILSAEHDGKLTLLTVDPAV 660
QY 661 GSVIG 665
DB 661 GSVIG 665

RESULT 3
US-09-432-695-2
; Sequence 2, Application US/09432695
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: Novel Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/432,695
; FILING DATE: 03-Nov-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,059
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 665 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-432-695-2

Query Match 100.0%; Score 3479; DB 18; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.8e-304;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSEKNFYITPIYVPSKGLHIGSAYTIIACDVLARYKRLMGYDVFYLTGLDEHGQKIQOK 60
DB 1 MSEKNFYITPIYVPSKGLHIGSAYTIIACDVLARYKRLMGYDVFYLTGLDEHGQKIQOK 60
QY 61 AEEAGITPQAVYDGMVGVKELWQLLDISYDKFIRTTDDHEKVVAAQVFERLLAQDDIYL 120
DB 61 AEEAGITPQAVYDGMVGVKELWQLLDISYDKFIRTTDDHEKVVAAQVFERLLAQDDIYL 120
QY 121 GEYSGWTSVSDDEEFTESQSLAEVFRDAGVNTGGIAPSGHEVEWVSESYFLRLSKYQDR 180
DB 121 GEYSGWTSVSDDEEFTESQSLAEVFRDAGVNTGGIAPSGHEVEWVSESYFLRLSKYQDR 180
QY 181 LVEFFKAHPEFITPDGRNLNEMLRNFIEPGLDLAVSRITTTGWVPVPSNPKHVVYVWIDA 240
DB 181 LVEFFKAHPEFITPDGRNLNEMLRNFIEPGLDLAVSRITTTGWVPVPSNPKHVVYVWIDA 240
QY 241 LLNTATAGYAQDEHGNFDFKFWNGTVFHMVCKOILRHSIYWPILLXWLDVKLPDRLIAH 300
DB 241 LLNTATAGYAQDEHGNFDFKFWNGTVFHMVCKOILRHSIYWPILLXWLDVKLPDRLIAH 300
QY 301 GWFVMDCKMSKSGNVVYPEMLVERYGLEDPLRYLMLNLPVSGSDGTFPTEDYVGRINYE 360
DB 301 GWFVMDCKMSKSGNVVYPEMLVERYGLEDPLRYLMLNLPVSGSDGTFPTEDYVGRINYE 360
QY 361 LANDLGNLNRVTSMINKYFDGOIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDYP 420
DB 361 LANDLGNLNRVTSMINKYFDGOIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDYP 420
QY 421 ALEAVWTLISRTNKYIDETAPWLDKDEALRDQASVMSHWQASIRVVAHLIEPFMMETS 480
DB 421 ALEAVWTLISRTNKYIDETAPWLDKDEALRDQASVMSHWQASIRVVAHLIEPFMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPFPRLNMBEETAYIKEQMEGNKPAV 540
DB 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPFPRLNMBEETAYIKEQMEGNKPAV 540
QY 541 EKWNPDVELKLNKDEIKFEDPKVEIRVAEVEKSVKVEGSDKLIQFLDAGDGEDRQI 600
DB 541 EKWNPDVELKLNKDEIKFEDPKVEIRVAEVEKSVKVEGSDKLIQFLDAGDGEDRQI 600
QY 601 LSGIAKYYPNEQELVKKVQIVANLKPRKMMKVVQSGMILSAEHDGKLTLLTVDPAV 660
DB 601 LSGIAKYYPNEQELVKKVQIVANLKPRKMMKVVQSGMILSAEHDGKLTLLTVDPAV 660
QY 661 GSVIG 665
DB 661 GSVIG 665

RESULT 4
US-10-606-607-2
; Sequence 2, Application US/10606607
; GENERAL INFORMATION:
; APPLICANT: LAWLOR, ELIZABETH JANE
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GM10253B-3CID1
; CURRENT APPLICATION NUMBER: US/10/606,607
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: 10/187,641
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 10/025,189
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/432,692
; PRIOR FILING DATE: 1999-11-03
; PRIOR APPLICATION NUMBER: 08/844,056
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: GB 9607999.1
; PRIOR FILING DATE: 1996-04-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665

; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-606-607-2

Query Match 100.0%; Score 3479; DB 32; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.8e-304;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSEKFPYITPTIYPYSGKLHIGSAITTIACDVLAARYKRLMGVDVFLYGLDEHGQKIQQK 60
DB 1 MSEKFPYITPTIYPYSGKLHIGSAITTIACDVLAARYKRLMGVDVFLYGLDEHGQKIQQK 60
QY 61 ABEAGITPQAYVDGMAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL 120
DB 61 ABEAGITPQAYVDGMAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL 120
QY 121 GEYSGWYSVDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSESYFRLLSKYQDR 180
DB 121 GEYSGWYSVDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSESYFRLLSKYQDR 180
QY 181 LVEFFKAHPEFITPDGRLNEMLRNFIPEGLEDLAVSRITFTWGVVPVSPNKHVVVWIDA 240
DB 181 LVEFFKAHPEFITPDGRLNEMLRNFIPEGLEDLAVSRITFTWGVVPVSPNKHVVVWIDA 240
QY 241 LLNATATGALGAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLIAH 300
DB 241 LLNATATGALGAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLIAH 300
QY 301 GFWFMKDGKMSKSGNVVPEMLVERYGDLPLRYILMRNLPGVSDGTFPTPEYVGRINYE 360
DB 301 GFWFMKDGKMSKSGNVVPEMLVERYGDLPLRYILMRNLPGVSDGTFPTPEYVGRINYE 360
QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSIADPHTHMEAVDYPR 420
DB 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSIADPHTHMEAVDYPR 420
QY 421 ALBAVWTLISRTNKYIDETAPVWLDKDEALRDLQASVMSHWAQSRVVAHLIEPPMMETS 480
DB 421 ALBAVWTLISRTNKYIDETAPVWLDKDEALRDLQASVMSHWAQSRVVAHLIEPPMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
DB 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
QY 541 EKWNPDVELKLNKDEIKFEDFDKVEIRVAEVEKSVKVEGSKLLQRLDAGDGEDROI 600
DB 541 EKWNPDVELKLNKDEIKFEDFDKVEIRVAEVEKSVKVEGSKLLQRLDAGDGEDROI 600
QY 601 LSGIAXYPNEQELVGKQVQIVANLKPROMKKYVSQGMILSAEHDKGLTLITVDPAVFN 660
DB 601 LSGIAXYPNEQELVGKQVQIVANLKPROMKKYVSQGMILSAEHDKGLTLITVDPAVFN 660
QY 661 GSVIG 665
DB 661 GSVIG 665

RESULT 5

US-10-472-928-1498
; Sequence 1498, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH
; FILE REFERENCE: P026925WO
; CURRENT FILING DATE: 2003-09-26
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 1498
; LENGTH: 665

; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: methionyl-tRNA synthetase (metG)
; OTHER INFORMATION: Cellular location: cytoplasm
; OTHER INFORMATION: Similar to strain R6 sequence 15902740 (0.E+01)
US-10-472-928-1498

Query Match 99.1%; Score 3446; DB 30; Length 665;
Best Local Similarity 99.2%; Pred. No. 1.7e-301;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSEKFPYITPTIYPYSGKLHIGSAITTIACDVLAARYKRLMGVDVFLYGLDEHGQKIQQK 60
DB 1 MSEKFPYITPTIYPYSGKLHIGSAITTIACDVLAARYKRLMGVDVFLYGLDEHGQKIQQK 60
QY 61 ABEAGITPQAYVDGMAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL 120
DB 61 ABEAGITPQAYVDGMAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL 120
QY 121 GEYSGWYSVDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSESYFRLLSKYQDR 180
DB 121 GEYSGWYSVDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSESYFRLLSKYQDR 180
QY 181 LVEFFKAHPEFITPDGRLNEMLRNFIPEGLEDLAVSRITFTWGVVPVSPNKHVVVWIDA 240
DB 181 LVEFFKAHPEFITPDGRLNEMLRNFIPEGLEDLAVSRITFTWGVVPVSPNKHVVVWIDA 240
QY 241 LLNATATGALGAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLIAH 300
DB 241 LLNATATGALGAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLIAH 300
QY 301 GFWFMKDGKMSKSGNVVPEMLVERYGDLPLRYILMRNLPGVSDGTFPTPEYVGRINYE 360
DB 301 GFWFMKDGKMSKSGNVVPEMLVERYGDLPLRYILMRNLPGVSDGTFPTPEYVGRINYE 360
QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSIADPHTHMEAVDYPR 420
DB 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSIADPHTHMEAVDYPR 420
QY 421 ALBAVWTLISRTNKYIDETAPVWLDKDEALRDLQASVMSHWAQSRVVAHLIEPPMMETS 480
DB 421 ALBAVWTLISRTNKYIDETAPVWLDKDEALRDLQASVMSHWAQSRVVAHLIEPPMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
DB 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
QY 541 EKWNPDVELKLNKDEIKFEDFDKVEIRVAEVEKSVKVEGSKLLQRLDAGDGEDROI 600
DB 541 EKWNPDVELKLNKDEIKFEDFDKVEIRVAEVEKSVKVEGSKLLQRLDAGDGEDROI 600
QY 601 LSGIAXYPNEQELVGKQVQIVANLKPROMKKYVSQGMILSAEHDKGLTLITVDPAVFN 660
DB 601 LSGIAXYPNEQELVGKQVQIVANLKPROMKKYVSQGMILSAEHDKGLTLITVDPAVFN 660
QY 661 GSVIG 665
DB 661 GSVIG 665

RESULT 6

PCT-US02-03987-13418
; Sequence 13418, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; FILE REFERENCE: ELITRA.028VPC
; CURRENT FILING DATE: 2002-02-02
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-09


```

; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 13418
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
PCT-US02-03987-13418

Query Match          99.1%; Score 3446; DB 1; Length 679;
Best Local Similarity 99.2%; Pred. No. 1.8e-301;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYTLGLDEHGQKIQQK 60
DB 15 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYTLGLDEHGQKIQQK 74
QY 61 AEEAGITPQAYVDGMAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 120
DB 75 AEEAGITPQAYVDGMAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 134
QY 121 GEYSWYVSDEBFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDR 180
DB 135 GEYSWYVSDEBFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDR 194
QY 181 LVEFFKAHPFITPDGRNLNMLNFIPLGLDLAVSRRTFTWGVVPSPKHHVYVWIDA 240
DB 195 LVEFFKAHPFITPDGRNLNMLNFIPLGLDLAVSRRTFTWGVVPSPKHHVYVWIDA 254
QY 241 LLNYATAGYAODEHGNFDFKFWNGTVFHMVGKDIIRPHSIYWPILLMMLDVKLPDLRIAH 300
DB 255 LLNYATAGYAODEHGNFDFKFWNGTVFHMVGKDIIRPHSIYWPILLMMLDVKLPDLRIAH 314
QY 301 GWFVKDGMKSKSGNVVYPEMLVERYGLDPLRYILMRNLPGVSDGTFTPEDYVGRINYE 360
DB 315 GWFVKDGMKSKSGNVVYPEMLVERYGLDPLRYILMRNLPGVSDGTFTPEDYVGRINYE 374
QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEPDHVLAEVAKSIADPHTHMEAVDPR 420
DB 375 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEPDHVLAEVAKSIADPHTHMEAVDPR 434
QY 421 ALEAVWTLISRTNKYIDETAPWVLKDEALRDLQALASVMSHQAASIRVVAHLIEPFFMETS 480
DB 435 ALEAVWTLISRTNKYIDETAPWVLKDEALRDLQALASVMSHQAASIRVVAHLIEPFFMETS 494
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPELNMEEBIEIAYKEQMEGNKPAV 540
DB 495 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPELNMEEBIEIAYKEQMEGNKPAV 554
QY 541 EKWNPDPEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVEGSDKLLQFLDAGDGEDROI 600
DB 555 EKWNPDPEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVEGSDKLLQFLDAGDGEDROI 614
QY 601 LSGIAKYYPNEQELVGKKVQIVANLKPRKMKKYSQGMILSAEHDGKLLTLITVDPAPVN 660
DB 675 GSVIG 679
;
; ORGANISM: Streptococcus pneumoniae
; US-09-815-242-13418

Query Match          99.1%; Score 3446; DB 23; Length 679;
Best Local Similarity 99.2%; Pred. No. 1.8e-301;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYTLGLDEHGQKIQQK 60
DB 15 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYTLGLDEHGQKIQQK 74
QY 61 AEEAGITPQAYVDGMAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 120
DB 75 AEEAGITPQAYVDGMAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 134
QY 121 GEYSWYVSDEBFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDR 180
DB 135 GEYSWYVSDEBFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDR 194
QY 181 LVEFFKAHPFITPDGRNLNMLNFIPLGLDLAVSRRTFTWGVVPSPKHHVYVWIDA 240
DB 195 LVEFFKAHPFITPDGRNLNMLNFIPLGLDLAVSRRTFTWGVVPSPKHHVYVWIDA 254
QY 241 LLNYATAGYAODEHGNFDFKFWNGTVFHMVGKDIIRPHSIYWPILLMMLDVKLPDLRIAH 300
DB 255 LLNYATAGYAODEHGNFDFKFWNGTVFHMVGKDIIRPHSIYWPILLMMLDVKLPDLRIAH 314
QY 301 GWFVKDGMKSKSGNVVYPEMLVERYGLDPLRYILMRNLPGVSDGTFTPEDYVGRINYE 360
DB 315 GWFVKDGMKSKSGNVVYPEMLVERYGLDPLRYILMRNLPGVSDGTFTPEDYVGRINYE 374
QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEPDHVLAEVAKSIADPHTHMEAVDPR 420
DB 375 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEPDHVLAEVAKSIADPHTHMEAVDPR 434
QY 421 ALEAVWTLISRTNKYIDETAPWVLKDEALRDLQALASVMSHQAASIRVVAHLIEPFFMETS 480
DB 435 ALEAVWTLISRTNKYIDETAPWVLKDEALRDLQALASVMSHQAASIRVVAHLIEPFFMETS 494
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPELNMEEBIEIAYKEQMEGNKPAV 540
DB 495 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPELNMEEBIEIAYKEQMEGNKPAV 554
QY 541 EKWNPDPEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVEGSDKLLQFLDAGDGEDROI 600
DB 555 EKWNPDPEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVEGSDKLLQFLDAGDGEDROI 614
QY 601 LSGIAKYYPNEQELVGKKVQIVANLKPRKMKKYSQGMILSAEHDGKLLTLITVDPAPVN 660
DB 675 GSVIG 679
;
; ORGANISM: Streptococcus pneumoniae
; US-09-815-242-13418

GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
```

Db 615 LSGIATYPNEQELVGVKKVQIVANLKPROMKKVYVSGQMILSAEHDGKLTLLTVDPAVFN 674
QY 661 GSVIG 665
Db 675 GSVIG 679

RESULT 8
US-10-072-851-13418
; Sequence 13418, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13418
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-072-851-13418

Query Match 99.1%; Score 3446; DB 26; Length 679;
Best Local Similarity 99.2%; Pred. No. 1.8e-301;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSEKNYITPTIYPSGKLHIGSAYTTIACDVLARYKRLMGVDVFLTGLDHHGQKIQQK 60
Db 15 MSEKNYITPTIYPSGKLHIGSAYTTIACDVLARYKRLMGVDVFLTGLDHHGQKIQQK 74
QY 61 ABEAGITPQAYVDMGVKELWQLDISYDKFIRTTDDYHEKVAQVFERLLAODDIYL 120
Db 75 ABEAGITPQAYVDMGVKELWQLDISYDKFIRTTDDYHEKVAQVFERLLAODDIYL 134
QY 121 GEYSGHYSDDEFFTESQLAEVRDEAGNTGGIAPSGHEVWVWSSESYFLRLSKYQDR 180
Db 135 GEYSGHYSDDEFFTESQLAEVRDEAGNTGGIAPSGHEVWVWSSESYFLRLSKYQDR 194
QY 181 LVEFFKAHPEFTTPDGLNEMLNFTPEGLDLAVSRITFTWGVPSNPKHVWVWIDA 240
Db 195 LVEFFKAHPEFTTPDGLNEMLNFTPEGLDLAVSRITFTWGVPSNPKHVWVWIDA 254
QY 241 LLNATYALGAQDEHGNFKDFWNGTVFHMVGKDILRFHSIYWPILLMLDVKLPRLIAH 300
Db 255 LLNATYALGAQDEHGNFKDFWNGTVFHMVGKDILRFHSIYWPILLMLDVKLPRLIAH 314
QY 301 GNFVMDGKXSKGNVYVPEMLVERYGIDPLRYILMRNLPGVSDGTFTPDYVGRINYE 360
Db 315 GNFVMDGKXSKGNVYVPEMLVERYGIDPLRYILMRNLPGVSDGTFTPDYVGRINYE 374
QY 361 LANDGLNLRNVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSIADPHTHMEAVDYPR 420
Db 375 LANDGLNLRNVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSIADPHTHMEAVDYPR 434
QY 421 ALEAVWTLISRTNKYIDETAPWVLKDDEALRDLQASVMSHWQASIRVVAHLIEPFWMTS 480

Db 435 ALEAVWTLISRTNKYIDETAPWVLKDDEALRDLQASVMSHWQASIRVVAHLIEPFWMTS 494
QY 481 RAVLTQIGLEVSLENLSLADFPADVTVVAKGTPFPRLNMBEETAYIKEQHEGKPAV 540
Db 495 RAVLTQIGLEVSLENLSLADFPADVTVVAKGTPFPRLNMBEETAYIKEQHEGKPAV 554
QY 541 EKWNPDVEVLKLNKDEIKFEDFKVEIRVAEYKSVKVEGSKLLQFRLDAGDGDROI 600
Db 555 EKWNPDVEVLKLNKDEIKFEDFKVEIRVAEYKSVKVEGSKLLQFRLDAGDGDROI 614
QY 601 LSGIATYPNEQELVGVKKVQIVANLKPROMKKVYVSGQMILSAEHDGKLTLLTVDPAVFN 660
Db 615 LSGIATYPNEQELVGVKKVQIVANLKPROMKKVYVSGQMILSAEHDGKLTLLTVDPAVFN 674
QY 661 GSVIG 665
Db 675 GSVIG 679

RESULT 9
US-10-282-122A-73893
; Sequence 73893, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73893
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73893

Query Match 99.1%; Score 3446; DB 28; Length 679;
Best Local Similarity 99.2%; Pred. No. 1.8e-301;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 MSEKNFYITPIYPSGKHLIGSAYTTIACDLVARYKRLMGYDVFLYTLGLDHEGQKIQQ 60
Db 15 MSEKNFYITPIYPSGKHLIGSAYTTIACDLVARYKRLMGYDVFLYTLGLDHEGQKIQQ 74
QY 61 AEEAGITPOAYVDGMVGVKELWQLLDISYDKFIRTTDDYHEKVVAAQVFERLLAQDDIYL 120
Db 75 AEEAGITPOAYVDGMVGVKELWQLLDISYDKFIRTTDDYHEKVVAAQVFERLLAQDDIYL 134
QY 121 GEYSWMYSVSDDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFRLSKYQDR 180
Db 135 GEYSWMYSVSDDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFRLSKYQDR 194
QY 181 LVEFFKAHPERTTPDGLNEMLNFIPEGLEDLAVSRITFTWGPVPSNPKHVVYVWIDA 240
Db 195 LVEFFKAHPERTTPDGLNEMLNFIPEGLEDLAVSRITFTWGPVPSNPKHVVYVWIDA 254
QY 241 LLNAYATGALGYAQDEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMMLDVKLPDRLLIAH 300
Db 255 LLNAYATGALGYAQDEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMMLDVKLPDRLLIAH 314
QY 301 GWFVMDGKMSKSGNVTYVPEMLVERYGDLPLRYLLMRNLVPGSDGTFTPEDYVGRINYE 360
Db 315 GWFVMDGKMSKSGNVTYVPEMLVERYGDLPLRYLLMRNLVPGSDGTFTPEDYVGRINYE 374
QY 361 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTDFHVLAEVAKSIADFTHEAVDYPR 420
Db 375 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTDFHVLAEVAKSIADFTHEAVDYPR 434
QY 421 ALEAVWTLISRTNKYIDETAPWVLKDDEALRDQLASVMSHWQASIRVVAHLIEPFMMETS 480
Db 435 ALEAVWTLISRTNKYIDETAPWVLKDDEALRDQLASVMSHWQASIRVVAHLIEPFMMETS 494
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIPPRLNMEEBIAIYKEQMEGNKPAV 540
Db 495 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIPPRLNMEEBIAIYKEQMEGNKPAV 554
QY 541 EKWNPDVEVLKLNKDEIKFEDPKVEIRVAEVKEVSKVSGDKLLQFRLDAGDGEDROI 600
Db 555 EKWNPDVEVLKLNKDEIKFEDPKVEIRVAEVKEVSKVSGDKLLQFRLDAGDGEDROI 614
QY 601 LSGIAKYYPNEQELVGKVKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVFN 660
Db 615 LSGIAKYYPNEQELVGKVKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVFN 674
QY 661 GSVIG 665
Db 675 GSVIG 679
```

```
RESULT 10
US-10-640-833-2751
; Sequence 2751, Application US/10640833
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; APPLICANT: Zeng, Qiondong
; APPLICANT: Opperman, Timothy
; APPLICANT: Houseweart, Chad Eric
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH03-15
; CURRENT APPLICATION NUMBER: US/10/640,833
; CURRENT FILING DATE: 2003-08-14
; PRIOR APPLICATION NUMBER: US 09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2751
```

```
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-640-833-2751
Query Match 98.6%; Score 3429; DB 32; Length 679;
Best Local Similarity 98.8%; Pred. No. 6.2e-300;
Matches 657; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 1 MSEKNFYITPIYPSGKHLIGSAYTTIACDLVARYKRLMGYDVFLYTLGLDHEGQKIQQ 60
Db 15 MSEKNFYITPIYPSGKHLIGSAYTTIACDLVARYKRLMGYDVFLYTLGLDHEGQKIQQ 74
QY 61 AEEAGITPOAYVDGMVGVKELWQLLDISYDKFIRTTDDYHEKVVAAQVFERLLAQDDIYL 120
Db 75 AEEAGITPOAYVDGMVGVKELWQLLDISYDKFIRTTDDYHEKVVAAQVFERLLAQDDIYL 134
QY 121 GEYSWMYSVSDDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFRLSKYQDR 180
Db 135 GEYSWMYSVSDDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFRLSKYQDR 194
QY 181 LVEFFKAHPERTTPDGLNEMLNFIPEGLEDLAVSRITFTWGPVPSNPKHVVYVWIDA 240
Db 195 LVEFFKAHPERTTPDGLNEMLNFIPEGLEDLAVSRITFTWGPVPSNPKHVVYVWIDA 254
QY 241 LLNAYATGALGYAQDEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMMLDVKLPDRLLIAH 300
Db 255 LLNAYATGALGYAQDEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMMLDVKLPDRLLIAH 314
QY 301 GWFVMDGKMSKSGNVTYVPEMLVERYGDLPLRYLLMRNLVPGSDGTFTPEDYVGRINYE 360
Db 315 GWFVMDGKMSKSGNVTYVPEMLVERYGDLPLRYLLMRNLVPGSDGTFTPEDYVGRINYE 374
QY 361 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTDFHVLAEVAKSIADFTHEAVDYPR 420
Db 375 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTDFHVLAEVAKSIADFTHEAVDYPR 434
QY 421 ALEAVWTLISRTNKYIDETAPWVLKDDEALRDQLASVMSHWQASIRVVAHLIEPFMMETS 480
Db 435 ALEAVWTLISRTNKYIDETAPWVLKDDEALRDQLASVMSHWQASIRVVAHLIEPFMMETS 494
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIPPRLNMEEBIAIYKEQMEGNKPAV 540
Db 495 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIPPRLNMEEBIAIYKEQMEGNKPAV 554
QY 541 EKWNPDVEVLKLNKDEIKFEDPKVEIRVAEVKEVSKVSGDKLLQFRLDAGDGEDROI 600
Db 555 EKWNPDVEVLKLNKDEIKFEDPKVEIRVAEVKEVSKVSGDKLLQFRLDAGDGEDROI 614
QY 601 LSGIAKYYPNEQELVGKVKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVFN 660
Db 615 LSGIAKYYPNEQELVGKVKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVFN 674
QY 661 GSVIG 665
Db 675 GSVIG 679
```

```
RESULT 11
US-10-282-122A-74415
; Sequence 74415, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Jianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
```

```
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUIPA.0344
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 74415
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
; US-10-282-122A-74415

Query Match      82.9%; Score 2883.5; DB 28; Length 666;
Best Local Similarity 81.9%; Pred. No. 1.3e-250;
Matches 546; Conservative 54; Mismatches 64; Indels 3; Gaps 2;

QY 1 MSEKNFYITPIYPYSGKLHIGSAITTIACDVLARYKRLMGYDVFLYTLGLDEHGQKIQQK 60
DB 1 MMKKPFVITPIYPYSGKLHIGSAITTIACDVLARYKRLMGHEVFLYTLGLDEHGQKIQQK 60
QY 61 AEEAGITPQAYVDMGAVGVKELMQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL 120
DB 61 AKEAGITPQYVDMNAKVKALMQLLDISYDTFIRTTDDYHEEVAVAAFEKLLAQQDIYL 120
QY 121 GEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDR 180
DB 121 GEYSGWYSVDSDEFFTESQLKEVFRDEGQVIGGIAPSGHEVWVSEESYFLRLSKYQDR 180
QY 181 LVFEFFKAHPFETPDGRNLNEMLNFTPEGLDLAVSRITFTWGVVPSPKHHVYVWIDA 240
DB 181 LVAFFKERPDPIQDGRNMENKXNFTEPGLDLAVSRITFTWGVVPSPDEKHVVYVWIDA 240
QY 241 LLNYATGALGYQAQDEHGNFDFKFWNGTVFHMVGKDIILRPHSIYWPILLMLLDVLPRLIAH 300
DB 241 LLNYATGALGYRQANHANFKFWNGTVFHMVGKDIILRPHSIYWPILLMLLDLPMPRLIAH 300
QY 301 GFWFMKDGKMSKGNVVPYEMLVVERGLDPLRYILMRSLPVGSDGTFPTPDYVGRINYE 360
DB 301 GFWFMKDGKMSKGNVVPYEMLVVERGLDPLRYILMRSLPVGSDGTFPTPDYVGRINYE 360
QY 361 LANDLGNLLNRTVSMINKYFDGOIPAYVEGVTDFHVLAEVAKSTADPHTHWEAVDYPR 420
DB 361 LANDLGNLLNRTVAMINKYFDGTPAYVONGTAFADLSDLAQADLAHVKHMEAVDYPR 420
QY 421 ALSEAVWTLISRTNKYIDETAPWYLDKDEALRDQLASVMSHWQASIRVVAHLIEPFMMETS 480
DB 421 ALSEAVWTLIARTNKYIDETAPWYLDKDEALRDQLASVMAHLAASRLVAHVLIQPFMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADPADVTYVAKGTPIFFPLNNEEBIAIYKEQVEGNKPAY 540
DB 481 AATMAQUGLEPVSJLSTLALDPAFPANTKVAVKGTPIFFPLDMEABIDYIKAQM-GDSSAI 539
```

```
QY 541 --EKWNPDEVELKLNKDEIKFEDFDKVEIRVAEVEKSVKESGDKLLQRLDAGGSDR 598
DB 540 SOEKWNPDEVELKLNKDEIKFEDFDKVEIRVAEVEKSVKESGDKLLQRLDAGGSDR 599
QY 599 QILSGIAKYPNEOELVGKQVIVANLKPRKMKKVVYSGMILSAEHDGKLTLLTVDPAY 658
DB 600 QILSGIAKYPNEOELVGKQVIVANLKPRKMKKVVYSGMILSAEHDGKLTLLTVDPAY 659
QY 659 PNGSVIG 665
DB 660 PNGSIIG 666

RESULT 12
US-10-415-182A-2228
; Sequence 2228, Application US/10415182A
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/415,182A
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12024
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 2228
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
; US-10-415-182A-2228
```

```
Query Match      82.9%; Score 2883.5; DB 30; Length 666;
Best Local Similarity 81.9%; Pred. No. 1.3e-250;
Matches 546; Conservative 54; Mismatches 64; Indels 3; Gaps 2;

QY 1 MSEKNFYITPIYPYSGKLHIGSAITTIACDVLARYKRLMGYDVFLYTLGLDEHGQKIQQK 60
DB 1 MMKKPFVITPIYPYSGKLHIGSAITTIACDVLARYKRLMGHEVFLYTLGLDEHGQKIQQK 60
QY 61 AEEAGITPQAYVDMGAVGVKELMQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL 120
DB 61 AKEAGITPQYVDMNAKVKALMQLLDISYDTFIRTTDDYHEEVAVAAFEKLLAQQDIYL 120
QY 121 GEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDR 180
DB 121 GEYSGWYSVDSDEFFTESQLKEVFRDEGQVIGGIAPSGHEVWVSEESYFLRLSKYQDR 180
QY 181 LVFEFFKAHPFETPDGRNLNEMLNFTPEGLDLAVSRITFTWGVVPSPKHHVYVWIDA 240
DB 181 LVAFFKERPDPIQDGRNMENKXNFTEPGLDLAVSRITFTWGVVPSPDEKHVVYVWIDA 240
QY 241 LLNYATGALGYQAQDEHGNFDFKFWNGTVFHMVGKDIILRPHSIYWPILLMLLDVLPRLIAH 300
DB 241 LLNYATGALGYRQANHANFKFWNGTVFHMVGKDIILRPHSIYWPILLMLLDLPMPRLIAH 300
QY 301 GFWFMKDGKMSKGNVVPYEMLVVERGLDPLRYILMRSLPVGSDGTFPTPDYVGRINYE 360
DB 301 GFWFMKDGKMSKGNVVPYEMLVVERGLDPLRYILMRSLPVGSDGTFPTPDYVGRINYE 360
QY 361 LANDLGNLLNRTVSMINKYFDGOIPAYVEGVTDFHVLAEVAKSTADPHTHWEAVDYPR 420
DB 361 LANDLGNLLNRTVAMINKYFDGTPAYVONGTAFADLSDLAQADLAHVKHMEAVDYPR 420
QY 421 ALSEAVWTLISRTNKYIDETAPWYLDKDEALRDQLASVMSHWQASIRVVAHLIEPFMMETS 480
DB 421 ALSEAVWTLIARTNKYIDETAPWYLDKDEALRDQLASVMAHLAASRLVAHVLIQPFMMETS 480
```

QY 481 RAVLTGILEVSSLESLADPPADVTVAAGTPIFFPRLNMEETIAYIKEQMEGNKPAV 540
Db 481 AATMAQLGLEPVSDELSTALADFPANTKVAKGTPFFRLDMEAEIDY:KAQM-GDSSAI 539
QY 541 --EKENWPEVEIKLNDKDEIKPDPKVEIRVAEVEKSVKESDKLLOFRLDAGGDR 598
Db 540 SQSEKWPPEEVALKSEKDVITFTFPAVEIRVAEVEKSVKESKLLRFRYDAGGQDR 599
QY 599 QILSGIAKYPNEQELVGKQVQIVANLKPRKMMKYSQGMILSABHGDKLTLTVDPAV 658
Db 600 QILSGIAKYPNEQELVGKQVQIVANLKPRKMMKYSQGMILSABHGDKLTLTVDSSV 659
QY 659 PNGSVIG 665
Db 660 PNGSIIG 666

RESULT 13
US-10-282-122A-72423
; Sequence 72423, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 72423
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-282-122A-72423

Query Match 82.9%; Score 2883; DB 28; Length 669;
Best Local Similarity 82.4%; Pred. No. 1.5e-250;
Matches 553; Conservative 46; Mismatches 64; Indels 8; Gaps 5;
QY 1 MSEKN-FVITTPYVPSGLHIGSAYTTIACDVLARYKRLMGVDPVLTG:DEHGQKIQ 59
Db 1 MTEQPFYITTPYVPSGLHIGSAYTTIACDVLARYKRLMGVDPVLTG:DEHGQKIQ 60

QY 60 KAEBAGITPOAYVDMAGVGVKELQQLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIY 119
Db 61 KSEAGITPOAYVDMAGVGVKELQQLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIY 120
QY 120 LGEVSGWYSVDEFFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQD 179
Db 121 LGEVSGWYSVDEFFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQD 180
QY 180 RLVEFFKAHPFITPDGRNLNEMLRNFIEPGLDELAVSRTTFTWGPVPSPKPHVYVWID 239
Db 181 RLVEFFKAHPFITPDGRNLNEMLRNFIEPGLDELAVSRTTFTWGPVPSPKPHVYVWID 240
QY 240 ALLNYATAGYQDAHAFNDFKFWNGTVFHMVWGDILRFHSHYWFILLMMLDLKPERLIA 299
Db 241 ALLNYATAGYQDAHAFNDFKFWNGTVFHMVWGDILRFHSHYWFILLMMLDLKPERLIA 300
QY 300 HGFVWMDGKMSKSGNVVYPEMLVERYGDLPLRYLMRNLPVSGDGTFTPEDYVGRINY 359
Db 301 HGFVWMDGKMSKSGNVVYPEMLVERYGDLPLRYLMRNLPVSGDGTFTPEDYVGRINY 360
QY 360 ELANDLGNLNRYSMINKYFDGQIPAYVEGTFDHFVLAEVAESKSIADFTTHMEAVDYP 419
Db 361 ELANDLGNLNRYSMINKYFDGQIPAYVEGTFDHFVLAEVAESKSIADFTTHMEAVDYP 420
QY 420 RALEAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHWQASIRVAHLIEPFMMET 479
Db 421 RALEAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHWQASIRVAHLIEPFMMET 480
QY 480 SRAVLTOIGLEBVSLENLSLADFPADVTVAAGTPIFFPRLNMEETIAYIKEQMEGNKPA 539
Db 481 SNAIMEQLGLAQGFLEGLAGLPTGVKVIAGTPIFFRLDMEAEIAYIKEQM-GDNSA 539
QY 540 V----EKNWNPDEVELKLNKDEIKPDEDFKVEIRVAEVEKSVKESDKLLOFRLDAGD 595
Db 540 IAAEEKEWNPDEVELKLNKDEIKPDEDFKVEIRVAEVEKSVKESDKLLOFRLDAGD 599
QY 596 EDRQILSGIAKYPNEQELVGKQVQIVANLKPRKMMKYSQGMILSABHGDKLTLTIV 654
Db 600 EDRQILSGIAKYPNEQELVGKQVQIVANLKPRKMMKYSQGMILSABHGDKLTLTIV 658
QY 655 DPAVNGSVIG 665
Db 659 DSSVNGSVIG 669

RESULT 14
US-10-415-182A-10044
; Sequence 10044, Application US/10415182A
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/415,182A
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12024
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 10044
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-10-415-182A-10044

Query Match 78.5%; Score 2732.5; DB 30; Length 665;
Best Local Similarity 78.2%; Pred. No. 6e-237;
Matches 519; Conservative 61; Mismatches 81; Indels 3; Gaps 3;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2004, 21:59:37 ; Search time 27 Seconds
(without alignments)
1110.332 Million cell updates/sec

Title: US-10-606-607-2

Perfect score: 3479
Sequence: 1 MSEKNFYITPTIYPSGKLH.....DGKLTLLTVDPVNGSVIG 665

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 16898 seqs, 45081143 residues

Total number of hits satisfying chosen parameters: 16898

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pcp.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pcp.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pcp.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pcp.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pcp.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pcp.*
- 7: /cgn2_6/ptodata/2/paa/US11_NEW_COMB.pcp.*
- 8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3446	99.1	679	1 PCT-US02-09107B-73893	Sequence 73893, A
2	2883.5	82.9	666	1 PCT-US02-09107B-74415	Sequence 74415, A
3	2883	82.9	669	1 PCT-US02-09107B-72423	Sequence 72423, A
4	2295	66.0	669	1 PCT-US02-09107B-75809	Sequence 75809, A
5	2288	65.8	677	1 PCT-US02-09107B-42475	Sequence 42475, A
6	2113	60.7	664	1 PCT-US02-09107B-60385	Sequence 60385, A
7	2013.5	57.9	658	8 US-60-622-712-354	Sequence 354, App
8	1949	56.0	659	1 PCT-US02-09107B-71524	Sequence 71524, A
9	1935.5	55.6	669	1 PCT-US02-09107B-46123	Sequence 46123, A
10	1931	55.5	657	1 PCT-US02-09107B-43933	Sequence 43933, A
11	1922.5	55.3	656	1 PCT-US02-09107B-71042	Sequence 71042, A
12	1911	54.9	662	1 PCT-US04-37204-4240	Sequence 4240, Ap
13	1678	48.2	674	1 PCT-US02-09107B-52642	Sequence 52642, A
14	1640.5	47.2	644	1 PCT-US02-09107B-51647	Sequence 51647, A
15	1572	45.2	645	1 PCT-US02-09107B-53102	Sequence 53102, A
16	1383	39.8	670	1 PCT-US02-09107B-46354	Sequence 46354, A
17	1180.5	34.2	651	1 PCT-US03-31318-11428	Sequence 11428, A
18	1189.5	34.2	628	1 PCT-US02-09107B-54439	Sequence 54439, A
19	1153.5	33.2	521	1 PCT-US02-09107B-76747	Sequence 76747, A
20	1148.5	33.0	650	1 PCT-US02-09107B-58709	Sequence 58709, A
21	1045	30.0	519	1 PCT-US02-09107B-62338	Sequence 62338, A
22	1042.5	30.0	519	1 PCT-US02-09107B-62166	Sequence 62166, A
23	1040	29.9	537	1 PCT-US02-09107B-63675	Sequence 63675, A
24	1037	29.8	519	1 PCT-US02-09107B-64482	Sequence 64482, A
25	871.5	25.1	512	1 PCT-US02-09107B-64107	Sequence 64107, A

ALIGNMENTS

RESULT 1

PCT-US02-09107B-73893
; Sequence 73893, Application PC/TUS0209107B
; GENERAL INFORMATION:

; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA_034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B

; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242

; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993

; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923

; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851

; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: 60/362,699

; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73893

; LENGTH: 679
; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae
PCT-US02-09107B-73893

Query Match 99.1%; Score 3446; DB 1; Length 679;
Best Local Similarity 99.2%; Pred. No. 1.1e-178; Indels 0; Gaps 0;
Matches 660; Conservative 2; Mismatches 3;

QY	1	MSEKNFYITPTIYPSGKLHIGSAVTTIACDVLARYKXLMGYDVFYLTGLDHEGQKIQQK	60
DB	15	MSEKNFYITPTIYPSGKLHIGSAVTTIACDVLARYKXLMGYDVFYLTGLDHEGQKIQQK	74
QY	61	AEAGITPOAYDGMAGVKGELWQLLDISYDKFIETDDYHEKVAQVFERLLAODDIYL	120
DB	75	AEAGITPOAYDGMAGVKGELWQLLDISYDKFIETDDYHEKVAQVFERLLAODDIYL	134
QY	121	GEYSGWYSVSDDEFFTESQAEVFRDEAGNVTGGIAPSGHEVEWSESYFLRLSKYQDR	180
DB	135	GEYSGWYSVSDDEFFTESQAEVFRDEAGNVTGGIAPSGHEVEWSESYFLRLSKYQDR	194
QY	181	LVEFFKAHPEFTTPDGRNLNEMLNFIETDGLDVAVSRTFTWGVPSNPKNVYVWIDA	240
DB	195	LVEFFKAHPEFTTPDGRNLNEMLNFIETDGLDVAVSRTFTWGVPSNPKNVYVWIDA	254
QY	241	LLNYATALGYAODEHGNFDFKFWNGVFMVWGKDLRFHSIYWPILLMMLDVKLPDLRIAH	300

26	838	24.1	512	1	PCT-US02-09107B-63402	Sequence 63402, A
27	755	21.7	675	1	PCT-US02-09107B-68671	Sequence 68671, A
28	729	21.0	677	1	PCT-US02-09107B-56600	Sequence 56600, A
29	729	21.0	677	1	PCT-US02-09107B-75322	Sequence 75322, A
30	729	21.0	677	1	PCT-US02-09107B-76244	Sequence 76244, A
31	726.5	20.9	593	6	US-10-930-328-3544	Sequence 9544, Ap
32	726	20.9	680	1	PCT-US02-09107B-60052	Sequence 60052, A
33	723.5	20.8	731	1	PCT-US02-09107B-77210	Sequence 77210, A
34	721.5	20.7	682	1	PCT-US02-09107B-66961	Sequence 66961, A
35	718	20.6	677	1	PCT-US02-09107B-55682	Sequence 55682, A
36	714.5	20.5	670	1	PCT-US02-09107B-61147	Sequence 61147, A
37	703	20.2	675	1	PCT-US02-09107B-78246	Sequence 78246, A
38	687	19.7	684	1	PCT-US02-09107B-65466	Sequence 65466, A
39	682.5	19.6	679	1	PCT-US02-09107B-48489	Sequence 48489, A
40	677.5	19.5	682	1	PCT-US02-09107B-58468	Sequence 58468, A
41	674.5	19.4	685	1	PCT-US02-09107B-65671	Sequence 65671, A
42	660.5	19.0	700	1	PCT-US02-09107B-51105	Sequence 51105, A
43	660	19.0	690	1	PCT-US02-09107B-44568	Sequence 44568, A
44	651	18.7	679	1	PCT-US02-09107B-67599	Sequence 67599, A
45	633.5	18.2	683	1	PCT-US02-09107B-63291	Sequence 63291, A

Db 255 LNNYATAGYAOQDEHGFDFKFWNGTVFHVGVKDLIRFHSIYWPILLMMLDVKLPDRLIAH 314
QY 301 GWFVKDGMKSKGNVVTPEMLVERYGDLDPRIYILMRNLPVGSDGTFPTPEYVGRINYE 360
Db 315 GWFVKDGMKSKGNVVTPEMLVERYGDLDPRIYILMRNLPVGSDGTFPTPEYVGRINYE 374
QY 361 LANDLGNLNRVTSMINKYFDGOIPAYVEGVTEFDHVLAEVAEKSTADPHTHMEAVDPR 420
Db 375 LANDLGNLNRVTSMINKYFDGOIPAYVEGVTEFDHVLAEVAEQSIADPHTHMEAVDPR 434
QY 421 ALBAVWTLISRTNKYIDETAPWVLDKDEALRDQASVMSHWQASIRVVAHLIEPFFMETS 480
Db 435 ALBAVWTLISRTNKYIDETAPWVLDKDEALRDQASVMSHWQASIRVVAHLIEPFFMETS 494
QY 481 RAVLTQIGLEEVSSLENLSIADFPADVTVVAKGTRIPFPRLNMEETIAYIKQMEGNKPAV 540
Db 495 RAVLTQIGLEEVSSLENLSIADFPADVTVVAKGTRIPFPRLNMEETIAYIKQMEGNKPAV 554
QY 541 EKAWNPDVELKLNKDEIKFEDFKVEIRVAEVEKSKVEGSDKLLQPRLDAGDGEDRQI 600
Db 555 EKAWNPDVELKLNKDEIKFEDFKVEIRVAEVEKSKVEGSDKLLQPRLDAGDGEDRQI 614
QY 601 LSGIAKYYPNEQELVGKQVQIVANLXPRKMKKYVSGMILSAEHDGKLTLLTVDPAPVN 660
Db 615 LSGIAKYYPNEQELVGKQVQIVANLXPRKMKKYVSGMILSAEHDGKLTLLTVDPAPVN 674
QY 661 GSVIG 665
Db 675 GSVIG 679

RESULT 2

PCT-US02-09107B-74415
; Sequence 74415, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 74415
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
PCT-US02-09107B-74415

Query Match 82.9%; Score 2883.5; DB 1; Length 666;
Best Local Similarity 81.9%; Pred. No. 2.4e-148;
Matches 546; Conservative 54; Mismatches 64; Indels 3; Gaps 2;

QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYGLDGHGKIQK 60
Db 1 MWKQFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGHEVFLYGLDGHGKIQK 60
QY 61 AEEAGITPQAYVDMGAVGKELWQLLDISYDKPRTDDYHEKVAQVPERLLAODDIYL 120
Db 61 AEEAGITPQAYVDMGAVGKELWQLLDISYDKPRTDDYHEKVAQVPERLLAODDIYL 120
QY 121 GEYSGWYSVDEBFFTESQLAEVFRDAGNVTGGTAPSGHEVWVSSESYFLRLSKYDDR 180

Db 121 GEYSGWYSVDEBFFTESQLAEVFRDAGNVTGGTAPSGHEVWVSSESYFLRLSKYDDR 180
QY 181 LVPEFFKAHPFIIPDGRLEMLRNFTIEFGLEDLAVSRRTFTTGWGPVPSNPKHVYVYVWIDA 240
Db 181 LVFAFFKEPFIIPDGRNEMWKNFTIEFGLEDLAVSRRTFTTGWGPVPSNPKHVYVYVWIDA 240
QY 241 LLNATATGALGAQDEHGFDFKFWNGTVFHVGVKDLIRFHSIYWPILLMMLDVKLPDRLIAH 300
Db 241 LLNATATGALGAQDEHGFDFKFWNGTVFHVGVKDLIRFHSIYWPILLMMLDVKLPDRLIAH 300
QY 301 GWFVKDGMKSKGNVVTPEMLVERYGDLDPRIYILMRNLPVGSDGTFPTPEYVGRINYE 360
Db 301 GWFVKDGMKSKGNVVTPEMLVERYGDLDPRIYILMRNLPVGSDGTFPTPEYVGRINYE 360
QY 361 LANDLGNLNRVTSMINKYFDGOIPAYVEGVTEFDHVLAEVAEKSTADPHTHMEAVDPR 420
Db 361 LANDLGNLNRVTSMINKYFDGOIPAYVEGVTEFDHVLAEVAEKSTADPHTHMEAVDPR 420
QY 421 ALBAVWTLISRTNKYIDETAPWVLDKDEALRDQASVMSHWQASIRVVAHLIEPFFMETS 480
Db 421 ALBAVWTLISRTNKYIDETAPWVLDKDEALRDQASVMSHWQASIRVVAHLIEPFFMETS 480
QY 481 RAVLTQIGLEEVSSLENLSIADFPADVTVVAKGTRIPFPRLNMEETIAYIKQMEGNKPAV 540
Db 481 RAVLTQIGLEEVSSLENLSIADFPADVTVVAKGTRIPFPRLNMEETIAYIKQMEGNKPAV 540
QY 541 --EKAWNPDVELKLNKDEIKFEDFKVEIRVAEVEKSKVEGSDKLLQPRLDAGDGEDR 598
Db 540 SQEKWVPEEVALKSEKDVITFTFDVAEIRVAEVEKSKVEGSEKLLAFRVDAGDQDR 599
QY 599 QILSGIAKYYPNEQELVGKQVQIVANLXPRKMKKYVSGMILSAEHDGKLTLLTVDPAPV 658
Db 600 QILSGIAKYYPNEQELVGKQVQIVANLXPRKMKKYVSGMILSAEHDGKLTLLTVDPAPV 659
QY 659 PNGSVIG 665
Db 660 PNGSVIG 666

RESULT 3

PCT-US02-09107B-72423
; Sequence 72423, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 72423
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Streptococcus mutans
PCT-US02-09107B-72423

Query Match 82.9%; Score 2883; DB 1; Length 669;
Best Local Similarity 82.4%; Pred. No. 2.6e-148;
Matches 553; Conservative 46; Mismatches 64; Indels 8; Gaps 5;
QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYGLDGHGKIQK 59
Db 1 MTEKQFYITPIYPSGKLHIGSAVTTIACDVLARYKRMNHDVFLYGLDGHGKIQK 60

```
Query Match      66.0%; Score 2295; DB 1; Length 669;
Best Local Similarity 65.2%; Pred. No. 1.4e-116;
Matches 438; Conservative 93; Mismatches 129; Indels 12; Gaps 9;

QY 60 KAEAGITPOAYVDGMAGVKELWQLLDISYDKFIRTTDDYHKKVAQVPERLLAODDYY 119
DB 61 KSEAGITPOAYVDGMAGVKELWQLLDISYDKFIRTTDDYHKKVAQVPERLLAODDYY 120
QY 120 LGYSGWYSVDEEFFTESQLAVERDEAGNVGTGGIAPSGHEVWVSESYFRLRSKYOD 179
DB 121 LGYSGWYSVDEEFFTESQLEIFRDEDKVIGGIAPSGHEVWVSESYFRLRSKYAD 180
QY 180 RLVEFFKAHPEFITPDGRNLEMRNFIEPGLEDLAVSRRTFTTGWVPVPSNPKHVYVWID 239
DB 181 RLVDFFHAHDFIOPDGRNVEIKNFIEPGLEDLAVSRRTFTTGWVPVPSNPKHVYVWID 240
QY 240 ALLNATYALGYADGHEGNFDPKFWNGTVFHMVGKDIILRFHSIYWPILLMLMDVCLPDLRIA 299
DB 241 ALLNATYALGYADGHEGNFDPKFWNGTVFHMVGKDIILRFHSIYWPILLMLMDVCLPDLRIA 300
QY 300 HGFVMDKGMKSKGNVVPPEMLVERYGLDPLRYILMRNLPVSGDGTFTTPEYVGRINX 359
DB 301 HGFVMDKGMKSKGNVVPPEMLVERYGLDPLRYILMRNLPVSGDGTFTTPEYVGRINX 360
QY 360 ELANDIGNLNRTVSMINKYFDGQIPAYVEGVTEFDRHVLAEVAEKSIADPHTHMEAVDYP 419
DB 361 ELANDIGNLNRTVSMINKYFDGQIPAYVEGVTEFDRHVLAEVAEKSIADPHTHMEAVDYP 420
QY 420 RALEAVNTLISRTNKYIDETAPWLDKDEALRDQLASVMSHWQASIRVVAHLIEPFMMET 479
DB 421 RALEAVNTLISRTNKYIDETAPWLDKDEALRDQLASVMSHWQASIRVVAHLIEPFMMET 480
QY 480 SRAVLTOIGLEEVSSLENLSLADFPADVTYVAKGTPIFPPLNMBEEIAYIKEQM--EG-- 539
DB 481 SNAIMEQLGLAGQDLEGLAGLPTGKVIKGTPIFPPLNMBEEIAYIKEQM--EG-- 540
QY 540 V-----EKEWNPDEVELKLNKDETFKDFDKVEIRVAEVEKSVKVGSDKLLQRLDAGD 595
DB 540 IAOEEKEWNPDEVELKLNKDETFKDFDKVEIRVAEVEKSVKVGSDKLLQRLDAGD 599
QY 596 EDQRLSGIAKYYPNEQELVGKVKQIVANLKPRKMKKYYVSGMILSAEH--DGKLTLLTV 654
DB 600 EDQRLSGIAKYYPNEQELVGKVKQIVANLKPRKMKKYYVSGMILSAEH--DGKLTLLTV 658
QY 655 DPVAPNGSVIG 665
DB 659 DSSVPNGSQIG 669

RESULT 4
PCT-US02-09107B-57809
; Sequence 57809, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57809
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Enterococcus faecium
PCT-US02-09107B-57809
```

```
Query Match      66.0%; Score 2295; DB 1; Length 669;
Best Local Similarity 65.2%; Pred. No. 1.4e-116;
Matches 438; Conservative 93; Mismatches 129; Indels 12; Gaps 9;

QY 1 MSEKTFYITPIYVPSGKLHIGSAYTTIACDLVARYKRLMGYDVLYLTGLDEHGQKIQQ 59
DB 1 MAEKETFTYITPIYVPSGKLHIGSAYTTIACDLVARYKRLMGYDVLYLTGLDEHGQKIET 60
QY 60 KAEAGITPOAYVDGMAGVKELWQLLDISYDKFIRTTDDYHKKVAQVPERLLAODDYY 119
DB 61 KAEELGVKQPQYVDGMAADVKKLWKLTDISYDKFIRTTDDYHKKAAVQKIFRLLLEQGGDIY 120
QY 120 LCEYSGWYSVDEEFFTESQLAVERDEAGNVGTGGIAPSGHEVWVSESYFRLRSKYOD 179
DB 121 LCEYSGWYSVDEEFFTESQLEIFRDEDKVIGGIAPSGHEVWVSESYFRLRSKYAD 180
QY 180 RLVEFFKAHPEFITPDGRNLEMRNFIEPGLEDLAVSRRTFTTGWVPVPSNPKHVYVWID 239
DB 181 RLLEYEYHSHPEFIOPESSRNEMINNFIXPGLEDLAVSRRTFTTGWVPVPSNPKHVYVWID 240
QY 240 ALLNATYALGYADGHEGNFDPKFWNGTVFHMVGKDIILRFHSIYWPILLMLMDVCLPDLRIA 299
DB 241 ALUSNTIITGALGYSGDESELSFEKIWPADV--QMGKELVRPHTIYWPILMLNALDLPKPKVFG 299
QY 300 HGFVMDKGMKSKGNVVPPEMLVERYGLDPLRYILMRNLPVSGDGTFTTPEYVGRINX 359
DB 300 HGWLLMKDGMKSKGNVVPPEMLVERYGLDPLRYILMRNLPVSGDGTFTTPEYVGRINX 359
QY 360 ELANDIGNLNRTVSMINKYFDGQIPAYVEGVTEFDRHVLAEVAEKSIADPHTHMEAVDYP 419
DB 360 DIANDIGNLNRTVSMINKYFDGQIPAYVEGVTEFDRHVLAEVAEKSIADPHTHMEAVDYP 419
QY 420 RALEAVNTLISRTNKYIDETAPWLDKDEALRDQLASVMSHWQASIRVVAHLIEPFMMET 479
DB 420 TALAELIWLISRTNKYIDETAPWLDKDEALRDQLASVMSHWQASIRVVAHLIEPFMMET 479
QY 480 SRAVLTOIGLEEVSSLENLSLADFPADVTYVAKGTPIFPPLNMBEEIAYIKEQM--EG-- 535
DB 480 PEKIFNOLGSETNLEGLHFGPEPPSGTKVYVAKGTPIFPPLNMBEEIAYIKEQM--EG-- 539
QY 536 -NKPAVEKEWNPDEVELKLNKDETFKDFDKVEIRVAEVEKSVKVGSDKLLQRLDAG 593
DB 540 TNEDTV--KWDPEETELVSTKEKQIKFDVFEKVELKVAEIVNCKVEGADKLLQRLDAG 597
QY 594 DGEDRQLSGIAKYYPNEQELVGKVKQIVANLKPRKMKKYYVSGMILSAE--HDGKLTLL 652
DB 598 DSQDRQLSGIAEYFPDPSSELIGKVKQIVANLKPRK--MRGQISQGMILSAEAPDGLQVI 656
QY 653 TVDPAVNGSVI 664
DB 657 EAPKSPNGSEI 668

RESULT 5
PCT-US02-09107B-42475
; Sequence 42475, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
```

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 42475

; LENGTH: 677

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

PCT-US02-09107B-42475

Query Match 65.8%; Score 2288; DB 1; Length 677;

Best Local Similarity 63.9%; Pred. No. 3.4e-116;

Matches 428; Conservative 104; Mismatches 130; Indels 8; Gaps 7;

QY 1 MSEK-NFYITPIYPSGKLHIGSAVTIACDVLARYKELMGYDVFLYLTGLDEHGQKIQQ 59

DB 9 MSQKEIFYITPIYPSGKLHIGNSYTIACDAMARYKELMGDFVFLYLTGVDEHQKLEK 68

QY 60 KAEAGITPAQYDGMAGVQKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIY 119

DB 69 KAAELNVTPEYVDKMAADYQKWLKTLDISYDKFIRTTDDYHMAAVQOIFDLRVEQGIY 128

QY 120 LGEYSWYSVSDDEEFTESQLAEVFRDEAGNVTGGAPSGHEVEMVSESYFLRLSKYQD 179

DB 129 LGEYEWYSVSDDEEFTETQLAEVYRDDGNGVIGGKAPSGHEVELVKEESYFFRNSKYAD 188

QY 180 RLVEFFKAHPEFITPQGRNLNEMLNPIEPGLEDLAVSRITFTTWGVVPSNPKHVVYVMD 239

DB 189 RLVOYVEEHPFIQPSRKNEMLNPIEPGLEDLAVSRITFTTWGVVPSNPKHVVYVMD 248

QY 240 ALANYATAGYAQDEHGNDFKNGTVFHMVGKDLIRFHSIYWPILLMMLDVKLPORLIA 299

DB 249 ALSNYITAGYGEDDSLFQKYPANV-QVNGKEIVRFHTIYWPILLMMLDPLPKKVPF 307

QY 300 HGFVMDGKMSKGNVYVPEMLVERYGLDPLRYLLMNLNLPVSGDTFTPDYVGRINY 359

DB 308 HGWLLMDGKMSKGNVYVPEMLVERYGLDPLRYLLRAIPFGSDGVFTPDYVGRINY 367

QY 360 ELANDNLNRTVSMINKYFDGQIPAYVEGTEFDHVAEVAEKSIAADFHTHMEADV 419

DB 368 DLANDNLNRTVSMINKYFDGQIPAYVEGTEFDHVAEVAEKSIAADFHTHMEADV 427

QY 420 RALEAVWTLSITNKYIDETAPWLDKDEALRDQASVNSHWAQASIRVVAHLIEPFMMET 479

DB 428 TATAEITWLVSRANKYIDETAPWLVAKESKEKELESVNIHAEISURIVAILLQPMET 487

QY 480 SRAVLTOIGLE-EVSSLENLSIADFPADVTVVAKGPIPPRLMEEBETAIYKEQMEGNK 538

DB 488 PGKIFPQLGLDPTMNMENIHFEFFPTDVTTSKGTPIFPRLEIETEVYIQQKMSQSES 547

QY 539 AVEKE--WNPDEVELKND-EIKFEDFDKVEIRVAEVEKSVKSGSKLQFRLDAGD 595

DB 548 ATBEDIKWPEETLVSTKEQIKYDDFDKVELKAEVDCCKVKGADKLLQFRLDAGD 607

QY 596 EDQILSGIAKYYPNEQELVKKVQIVANLKPROMKKYVSGMILSAEH-DGKLTLLTV 654

DB 608 NHRQILSGIAEFPDPAALIGKVVIVANLKPROMKKYVSGMILSAEHPGKLIQIVEA 666

QY 655 DPAVNGSVI 664

DB 667 PKEMPNGAGI 676

RESULT 6

PCT-US02-09107B-60385

; Sequence 60385; Application PC/TUS0209107B

; GENERAL INFORMATION:

; APPLICANT: Elitra Pharmaceuticals Inc.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034VPC

; CURRENT APPLICATION NUMBER: PCT/US02/09107B

; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: 09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 09/948,993

; PRIOR FILING DATE: 2001-09-06

; PRIOR APPLICATION NUMBER: 60/342,923

; PRIOR FILING DATE: 2001-10-25

; PRIOR APPLICATION NUMBER: 10/072,851

; PRIOR FILING DATE: 2002-02-08

; PRIOR APPLICATION NUMBER: 60/362,699

; PRIOR FILING DATE: 2002-03-06

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 60385

; LENGTH: 664

; TYPE: PRT

; ORGANISM: Listeria monocytogenes

PCT-US02-09107B-60385

Query Match 60.7%; Score 2113; DB 1; Length 664;

Best Local Similarity 60.7%; Pred. No. 9.1e-107;

Matches 407; Conservative 101; Mismatches 144; Indels 18; Gaps 8;

QY 3 EKN-FYITPIYPSGKLHIGSAVTIACDVLARYKELMGYDVFLYLTGLDEHGQKIQQKA 61

DB 4 EKNTFYITPIYPSGKAHIGHAYTVAGDAMARYKELXGVDVFLYLTGTDHEGQKIQA 63

QY 62 BEAGITPAQYDGMAGVQKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIY 121

DB 64 KERGISEQYVDSIAEGFQELMKLEISNTDFIRTTQDRHKTSEVKIFEQLLEQGGIY 123

QY 122 EYSWYVSDEEFTESQLAEVFRDEAGNVTGGAPSGHEVEMVSESYFLRLSKYQDRL 181

DB 124 EYEGWYVSDEEFTETQLEEVYRDDGNGVIGGKAPSGHEVELVKEESYFFRNSKYADRL 183

QY 182 VSEFFKAHPEFITPQGRNLNEMLNPIEPGLEDLAVSRITFTTWGVVPSNPKHVVYVMD 241

DB 184 VEYNSHPEFIPESRKNEMLNPIEPGLEDLAVSRITFTTWGVVPSNPKHVVYVMD 243

QY 242 LNYATAGYAQDEHGNDFKNGTVFHMVGKDLIRFHSIYWPILLMMLDVKLPDLIAHG 301

DB 244 SNYITAGYNTDNTKPKYWPADV-QVNGKEIVRFHTIYWPILLMMLDPLPKVFGHG 302

QY 302 WFMVMDGKMSKGNVYVPEMLVERYGLDPLRYLLMNLNLPVSGDTFTPDYVGRINYL 361

DB 303 WILMDGKMSKGNVYVPEMLRIDRYLLREVLPFGSDGLFTPEDFVDRVNFOL 362

QY 362 ANDLGNLNTVSMINKYFDGQIPAYVEGTEFDHVAEVAEKSIAADFHTHMEADV 421

DB 363 ANDLGNLNTVSMINKYFDGQIPAYVEGTEFDHVAEVAEKSIAADFHTHMEADV 422

QY 422 LEAVWTLSITNKYIDETAPWLDKDEALRDQASVNSHWAQASIRVVAHLIEPFMMETSR 481

DB 423 LNQLWSLISRTNKYIDETAPWLVAKESKEKELESVNIHAEISURIVAILLQFLTRTPG 482

QY 482 AVTQIGLEEVSSLENL-----SLADFPADVTVVAKGPIPPRLMEEBETAIYKEQME 534

DB 483 EIFLQLQLOB-----ENLKKWDSIYGEIPAGTVVYKGTPIFPRLEDAEVETVYIODEMK 538

QY 535 GNPFAVEKWNPDVELKLNKDKIPEDPKVIRVAEVEKSVKSGSKLQFRLDAGD 594

DB 539 GSAPAPAE--VAEVE-ALETPOIGEDFDKIDIRVAEVEKQVQDKVKKADKLLCFQDLG 595

QY 595 GEDRQILSGIAKYYPNEQELVKKVQIVANLKPROMKKYVSGMILSAEHDKLTLLTV 654

DB 596 GKLRQVLSGIAEFPDPAALIGKVVIVANLKPROMKKYVSGMILSAEHPGKLSVIEA 653

QY 655 DPAVNGSVI 664

DB 654 SSALPNGAKV 663

RESULT 7

US-60-622-712-354

; Sequence 354; Application US/60622712

; GENERAL INFORMATION:

; APPLICANT: Klaenhammer, Todd

; APPLICANT: Russell, William


```

RESULT 9
PCT-US02-09107B-46123
; Sequence 46123, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46123
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US02-09107B-46123

Query Match      55.6%; Score 1935.5; DB 1; Length 669;
Best Local Similarity 57.0%; Pred. No. 3.5e-97;
Matches 383; Conservative 107; Mismatches 161; Indels 21; Gaps 12;

Qy 1 MSEKN--FYITTPYPSGKLHGSAITTIACDLVARYKRLMGYDVFYLTGLDEHGQIKI 58
Db 10 MTEENKSFYITTPYPSGKLHGSAITTVAGDAMARYKMGQYNNVYLTGTDEHGQIKI 69

Qy 59 QKAEAGITPOAYVDMGVKELWQLDIDSYDKFTRTDDYHEKVAQVFERLLAQQDI 118
Db 70 KKAELNVTPOAYVDMGVKELWQLDIDSYDKFTRTDDYHEKVAQVFERLLAQQDI 129

Qy 119 YLGEYSGWYSVDEEFFTSQLAQAEVFRDEAGNVGTGAP--SGHEVEWSEESYFRLSKY 177
Db 130 YLDEYSGWYSVDEEFFTSQLAQAEVFRDEAGNVGTGAP--EGDKVVGKSPDSGDHVELVREESYFRMGKY 188

Qy 178 QDLRVFFKAHPFITPDGRLNEMLNFIPEGLEDLAVSRITFTTGWGVPVSPNPKHVYVW 237
Db 189 VDRLLAFYEDNPHFIQPSRKNEMINNF.KPGLEDLAVSRITFTTGWGVPVSPNPKHVYVW 248

Qy 238 IDALLNATAGYAQDEHGNDFKFWNGTVFHMVGKDILRFHSIYWPILLMLDVLKLDL 297
Db 249 VDALSNTTALGVTGENEYKFWPADV--HLVKGKIVRFHSIYWPILLMALDPLPKKV 307

Qy 298 IAHGFWMDGKMSKGNVYVPEMLVERYGDLPLRYILMRNLPGVSDGTFPTEDYVGR 357
Db 308 FAHGWLMDGKMSKGNVDPVILIDRYGLDALRYLLREVPPGSDGVFTPEGFVERI 367

Qy 358 NYELANDLGNLNRVTSNMYKDFDQIPAYVEGVTEPFDHVLAEVAKSIADPHTHMEAVD 417
Db 368 NFEDLANDLGNLNRVTSNMYKDFDQIPAYVEGVTEPFDHVLAEVAKSIADPHTHMEAVD 427

Qy 418 YPRALAVNTLSRNTKYIDETAPWLDKDEALRDLQASVMHQAQSRVVAHLTEPFMM 477
Db 428 FSVALSISIQVLSRNTKYIDETAPWLDKDEALRDLQASVMHQAQSRVVAHLTEPFMM 487

Qy 478 ETSRAVLTOIGL--EVSLENLS--LADPPADVTVVAKGTPIPPRLNMBEETAYIKEQME 534
Db 488 VAPSKPAQGLTDEAHTSWGSLSTIGCIPAG--TKVEKGQPIFPRLMDVVEAYIKEQMK 546

Qy 535 GKNPAYE--KENPDEVELKLNKDEIKFDFDKVEIRVAEVKESVKGSDKLLQFLRDA 592
Db 547 ASAPKVEEKEEP-----KAEETIDDFKVELRVAEVLASAEFVKADKLLQIQLDL 599

Qy 593 GGGEDROIISGIAKYPNQELVKGVQIVANLKRPMKKYVSGMILSAEHDGKLTLL 652
Db 600 GT-EKRVQVSGIAKPY-SPEDLKGKVKVCTNVLKPVK-LRGELSQGMILAGEENGVLSLA 656

PCT-US02-09107B-43933
; Sequence 43933, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43933
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
PCT-US02-09107B-43933

Query Match      55.5%; Score 1931; DB 1; Length 657;
Best Local Similarity 56.3%; Pred. No. 6e-97;
Matches 377; Conservative 110; Mismatches 163; Indels 20; Gaps 11;

Qy 1 MSEKNFYITTPYPSGKLHGSAITTIACDLVARYKRLMGYDVFYLTGLDEHGQIQK 60
Db 1 MAKETFTTTPYPSGKLHGSAITTVAGDAMARYKMGQYDVRVLTGTDEHGQIQK 60

Qy 61 ABERAGITPOAYVDMGVKELWQLDIDSYDKFTRTDDYHEKVAQVFERLLAQQDIYL 120
Db 61 AQKAGTEIEYLDDEMIAGIKQLWAKLEISNDDEIRTEERKHVVEQVFERLLKQGDYIL 120

Qy 121 GEYSGWYSVDEEFFTSQLAQAEVFRDEAGNVGTGAP--SGHEVEWSEESYFRLSKYQD 179
Db 121 GEYEGWYSVDEEFFTSQLAQAEVFRDEAGNVGTGAP--QYENGKILGKSPDSGHEVELVREESYFFNISKYTD 179

Qy 180 RLVEFFKAHPFITPDGRLNEMLNFIPEGLEDLAVSRITFTTGWGVPVSPNPKHVYVWID 239
Db 180 RLLEFYDQNPDTFQPPSRKNEMINNF.KPGLEDLAVSRITFTTGWGVPVSPNPKHVYVWID 239

Qy 240 ALLNATAGYAQDEHGNDFKFWNGTVFHMVGKDILRFHSIYWPILLMLDVLKLDL 299
Db 240 ALVNYISALGYLSDDESLEFNKYWPADI--HLMAKEIVRFHSIYWPILLMALDPLPKKVFA 298

Qy 300 HGVFVMDGKMSKGNVYVPEMLVERYGDLPLRYILMRNLPGVSDGTFPTEDYVGRIN 359
Db 299 HGMILMDGKMSKGNVDPNLIIDRYGLDALRYLLREVPPGSDGVFTPEAFVETNF 358

Qy 360 ELANDLGNLNRVTSNMYKDFDQIPAYVEGVTEPFDHVLAEVAKSIADPHTHMEAVDP 419
Db 359 DLANDLGNLNRVTSNMYKDFDQIPAYVEGVTEPFDHVLAEVAKSIADPHTHMEAVDP 418

Qy 420 RALEAVNTLSRNTKYIDETAPWLDKDEALRDLQASVMHQAQSRVVAHLTEPFMMET 479
Db 419 VALSTVWKFSRNTKYIDETAPWLDKDEALRDLQASVMHQAQSRVVAHLTEPFMMET 478

Qy 480 SRAVLTOIGL-----EVSLENLS--LADPPADVTVVAKGTPIPPRLNMBEETAYIKEQME 534
Db 479 PKSIFEQNLNPNQFMEFSLEQYGV--LTESIMVTGQPKPIFPRLDSEATYIKESMQ 536

```

QY 535 GNKPAVEKWNPDVELKLNKDEIKFEDFKVEIRVAEVEKSVKSGSKLQFRDLADG 594
DB 537 --PPATSE--KEEIP---SKPQIDIKDFKVEIKAAITIDAHEVKSJLQVLDL-D 588
QY 595 GEDRQILSGIAKYYPNEQLVGGKVOIVANLKRKMMKKYVSGMILSAHDKLTLTV 654
DB 589 SEQRQIVSGIAKPY-TPDDIIGKKVAVVTLNPKAKLMGQ-KSEGMLSAEKGVLTVSL 646
QY 655 DPAVNGSVI 664
DB 647 PSAIPNGAVI 656

RESULT 11
PCT-US02-09107B-71042
; Sequence 71042, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78814
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 71042
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
PCT-US02-09107B-71042

Query Match 55.3%; Score 1922.5; DB 1; Length 666;
Best Local Similarity 55.9%; Pred. No. 1.7e-96;
Matches 375; Conservative 113; Mismatches 160; Indels 23; Gaps 11;

QY 1 MSEKNFYITPIYPSGKLHIGSAITTIACDVLARYKELMGYDVYLTGLDEHGKIQOK 60
DB 1 MAKDTFYITPIYPSGKLHIGSAITTVAGDVIAIRYKMQGYDVRYLTGTDEHGKIQOK 60

QY 61 ABEAGITPOAVDGMVGVKELWQLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYL 120
DB 61 AOKAGTELEVLDWISGINKLWSKLEISNDDFIRTEERHKQVVEKVERLLAQDDIYL 120

QY 121 GEYSGWYSVDEEFFTESQLAE-VFRDBAGNVGTGIAP-SGHEVWVSESEYFURLSKYQ 178
DB 121 GEYEGWYSPDETYYTESQLDVPVY--ENGKIVGKSPDSGHEVELKEESYFFNINKYT 178

QY 179 DRLVEFFKAEFEITPDGRINLEMLNFTPEGLDLAVSRRTFTTGWVPSNPKHVVYWI 238
DB 179 DRLLFEYDENPDFQPPSRKKNEMNFIKPGLEDLAVSRRTFSDGVRVPSNPKHVVYWI 238

QY 239 DALLNYATAGYAQDEHGNFDFKFWNGTVFHMVGKDILAFHSIYWPILMLMDVKLPRLI 298
DB 239 DALVNYISSGLYSDDETLENKYWPADI-HLMAKEIVRFHSIIPILIMALDLPKRVF 297

QY 299 AHGFWFKDGMKSKSGKNVVPPEMLVERYGLDPLRYILMRNLVPGSDGTFTPEYVGRIN 358
DB 298 AHGWTLMKDGKMSKSGKNVDPVNLIDRYGLDTRYYLMLRELPPGSDGVFTPEAFVERTN 357

QY 359 YELANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDDHVLAEVAKESIADFTHTMEAVDY 418
DB 358 YDLANDLGNLNRVTSMINKYFHGELPAYQGGKHELDKMEAMALETVKSFNDNMESLQF 417

QY 419 PRALEANVTLISRTNKYIDETAPVLDKDEALRDQALASVMHQASIRVVAHLIEPFWME 478
DB 418 SVALSTWKFISRTNKYIDETQPVWLAKDENQREMLGNVAHLVENIRFAILLQPFETH 477
QY 479 TSRAVLTIQIG-----LEEVSLENLSLADFPADVTVVAKGTPPIFPRLNMESEIAYIKEM 533
DB 478 APREIFKQNLINNPDLHOLDLSIQVGM--LSEAITVTEKPTPIFPRLDTEAEIAYIKESM 535
QY 534 EGNKPAVEKWNPDVELKLNKDEIKFEDFKVEIRVAEVEKSVKSGSKLQFRDLADG 593
DB 536 QPPK-SIKSDEP-----GREQIDIKDFKVEIKAAITIDAENVKSEKLLKIKVEL- 586
QY 594 DGEDRQILSGIAKYYPNEQLVGGKVOIVANLKRKMMKKYVSGMILSAHDKLTLTV 653
DB 587 DNEQRQIVSGIAKPY-RPEDIIGKKVAVVTLNPKAKLMGQ-KSEGMLSAEKGVLTVSL 644
QY 654 VDPVNGSVI 664
DB 645 LPSAIPNGAVI 655

RESULT 12
PCT-US04-37204-4240
; Sequence 4240, Application PC/TUS0437204
; GENERAL INFORMATION:
; APPLICANT: NOVOZYMES BIOTECH, INC.
; APPLICANT: NOVOZYMES A/S
; TITLE OF INVENTION: Bacillus Licheniformis Chromosome
; FILE REFERENCE: 10588.204-WO
; CURRENT APPLICATION NUMBER: PCT/US04/37204
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: 60/535,988
; PRIOR FILING DATE: 2004-01-09
; PRIOR APPLICATION NUMBER: 60/561,059
; PRIOR FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 60/572,403
; PRIOR FILING DATE: 2004-05-19
; NUMBER OF SEQ ID NOS: 8395
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4240
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
PCT-US04-37204-4240

Query Match 54.9%; Score 1911; DB 1; Length 662;
Best Local Similarity 56.4%; Pred. No. 7.2e-96;
Matches 375; Conservative 107; Mismatches 173; Indels 10; Gaps 8;

QY 3 EKNFYITPIYPSGKLHIGSAITTIACDVLARYKELMGYDVYLTGLDEHGKIQOKA 61
DB 4 EKNTFYITPIYPSGKLHIGSAITTVAGDMARYKRLGDFVRYLTGTDEHGKIQOFTA 63

QY 62 BEAGITPOAVDGMVGVKELWQLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYL 121
DB 64 EXENITPQELVDRAAEDIQOLWKLLDISNDDFIRTEERHKVIEKVFQKLLDNGDIYLD 123

QY 122 EYSGWYSVDEEFFTESQLAEVFRDBAGNVGTGIAP-SGHEVWVSESEYFURLSKYQOR 180
DB 124 EYEGWYSIPDETYYTETQLVDVERNEKGEVIGKSPDSGHPVELIKESEYFFRMGKYADR 183

QY 181 LVEFFKAEFEITPDGRINLEMLNFTPEGLDLAVSRRTFTTGWVPSNPKHVVYWI 240
DB 184 LLAFYENPEFTQPPSRKKNEMNFIKPGLEDLAVSRRTFDMGKIVGPNKPKHVVYWI 243

QY 241 LLNYATAGYAQDEHGNFDFKFWNGTVFHMVGKDILAFHSIYWPILMLMDVKLPDLIAH 300
DB 244 LFNYITALGENTENDENYQKYPADV-HLVGKEIVRFHSIIPILMALDLPKRVFAH 302

QY 301 GWFWMKDGKMSKSGKNVVPPEMLVERYGLDPLRYILMRNLVPGSDGTFTPEYVGRINYE 360
DB 303 GWTLMKDGKMSKSGKNVDPVNLIDRYGLDTRYYLMLRELPPGSDGVFTPEGEFVERINYD 362

QY 361 LANDGLNLRNTSMINKYFDGQIPAYVEGVTFBPDHVLAEVAEKSIADPHTHMEAVDYP 420
 DB 363 LANDGLNLRNTSMINKYFDGQIPAYVEGVTFBPDHVLAEVAEKSIADPHTHMEAVDYP 422
 QY 421 ALBAVWTLIRNTSMINKYFDGQIPAYVEGVTFBPDHVLAEVAEKSIADPHTHMEAVDYP 480
 DB 423 ALBAVWTLIRNTSMINKYFDGQIPAYVEGVTFBPDHVLAEVAEKSIADPHTHMEAVDYP 482
 QY 481 RAVTQIGLEEVSEVLENSLADPADVTY-VAKGTIPFRLNMEEEIAYI----KEQMEGN 539
 DB 483 EKIPAGLVGTDAKLTWDSQSGQLKSVTVQKGEPLFRLEAEDEVAVIKKMQGTAP- 541
 QY 540 VEKEMNDEVELKLNKDEIFEDFDKVEIRVAEVEKSVKVEGSKLLQFLDAGDGEDRQ 599
 DB 542 --KEEPQKEEKAHERLPEITIDDFMSTELRAVEIHAEPVKKADRLKQLDLG-PEKQ 598
 QY 600 ILSGIAKYYPNEBELVKKVQIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVP 659
 DB 599 VWSGIAKHYKPE-ELVGRKVICVTLNKPVK-LRGEISQGMILAGEDNGVLSAAVDSSLA 656
 QY 660 NGSVI 664
 DB 657 NGTRI 661

RESULT 13

PCT-US02-09107B-52642
 ; Sequence 52642, Application PC/TUS0209107B
 ; GENERAL INFORMATION:
 ; APPLICANT: Elittra Pharmaceuticals Inc.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034VPC
 ; CURRENT APPLICATION NUMBER: PCT/US02/09107B
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 09/815,242
 ; PRIOR FILING DATE: 2001-09-06
 ; PRIOR APPLICATION NUMBER: 09/948,993
 ; PRIOR FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: 60/342,923
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: 10/072,851
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: 60/362,699
 ; PRIOR FILING DATE: 2002-03-06
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 52642
 ; LENGTH: 674
 ; TYPE: PRT
 ; ORGANISM: Clostridium botulinum

PCT-US02-09107B-52642
 Query Match 48.2%; Score 1678; DB 1; Length 674;
 Best Local Similarity 48.7%; Pred. No. 2.8e-83;
 Matches 326; Conservative 136; Mismatches 176; Indels 32; Gaps 10;
 QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDFVLTGLDGHGKIQOK 60
 DB 30 MSKGTGYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDFVLTGLDGHGKIQOK 89
 QY 61 ABEAGITPOAYVDGMAGVKELMQLLDISYDKFIRTTDDYHEKVAQVFERLLAODDIY 120
 DB 90 ABAKVTTPKYVDYDEIVAGIKELMKNWNIYDKFIRTTDEYHVELVQKFKTYLDKGLYK 149
 QY 121 GEYSGWYSVDEEFFTESQLAEVFRDAGNVTGGIAPSGHEVWSESYFLRLSKYQDR 180
 DB 150 SEYSGWYCTPCSEFWTETQLED-----GKPCDCRSVEKAKEEAYFFKSKYAPK 199
 QY 181 LVEFFKAHPEFTPDGLRNLNENFTERGLDLANSRTTFTTGVDPVSPNPKHVVWIDA 240
 DB 200 LLEYIESHEPEFTQPEERKKNMNNFLKPLQDLQCVSRFSFDGIGIPVTFDEKRVIVW-DA 259
 QY 241 LLNYATALGYAQDEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMLMDVCLPDLIAH 300

DB 260 LANVITAGSYSTKDEDELYKKTWPDV-HLVKGDILRPHHTIYWPIMLMALDILPKQVFGH 318
 QY 301 GWFYMKDGKMSKSGNVVYPEMLVERYGLOPLRYLLMRNLFPVSGDGTFTPEDVYGRINYE 360
 DB 319 GLLVVDGKMSKSGNVVDPVVLVNHFGTDPVRYLLREIPFGADGIFNNEIFIKKINS 378
 QY 361 LANDGLNLRNTSMINKYFDGQIPAYVEGVTFBPDHVLAEVAEKSIADPHTHMEAVDYP 420
 DB 379 LANDGLNLRNTSMINKYFDGQIPAYVEGVTFBPDHVLAEVAEKSIADPHTHMEAVDYP 437
 QY 421 ALEAVWTLIRNTSMINKYFDGQIPAYVEGVTFBPDHVLAEVAEKSIADPHTHMEAVDYP 480
 DB 438 ALEAVWTLIRNTSMINKYFDGQIPAYVEGVTFBPDHVLAEVAEKSIADPHTHMEAVDYP 497
 QY 481 RAVTQIGLEEVSEVLENSLADPADVTY-VAKGTIPFRLNMEEEIAYI----KEQMEGN 536
 DB 498 IKINEQIN-ADVLNWDLSKFNKGTGKVKGEAIPRIDVEKKIBELNKLKEEQLKQ 556
 QY 537 KPAVEKSWNPDEVELKLNKDEIFEDFDKVEIRVAEVEKSVKVEGSKLLQFLDAGDGE 596
 DB 557 KAK-----NIAPIKEEITIDFEKIDLRVAKVLECEPIKGAKKLLKLVSLG-GE 605
 QY 597 DRQILSGIAKYYPNEBELVKKVQIVANLKPROMKKYVSQGMILSAEHDGKLTLLTV 654
 DB 606 ERQVVSIGTAKYKPE-DLIGKVVLVANLKPVK-LRGEISQGMILAASTDDSELSFVASI 663
 QY 655 DPAVNGSVI 664
 DB 664 DGDIEGNIV 673

RESULT 14

PCT-US02-09107B-51647
 ; Sequence 51647, Application PC/TUS0209107B
 ; GENERAL INFORMATION:
 ; APPLICANT: Elittra Pharmaceuticals Inc.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034VPC
 ; CURRENT APPLICATION NUMBER: PCT/US02/09107B
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 09/815,242
 ; PRIOR FILING DATE: 2001-09-06
 ; PRIOR APPLICATION NUMBER: 09/948,993
 ; PRIOR FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: 60/342,923
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: 10/072,851
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: 60/362,699
 ; PRIOR FILING DATE: 2002-03-06
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 51647
 ; LENGTH: 644
 ; TYPE: PRT
 ; ORGANISM: Clostridium acetobutylicum

PCT-US02-09107B-51647

Query Match 47.2%; Score 1640.5; DB 1; Length 644;
 Best Local Similarity 48.3%; Pred. No. 2.7e-81;
 Matches 322; Conservative 125; Mismatches 193; Indels 27; Gaps 10;
 QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDFVLTGLDGHGKIQOK 59
 DB 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDFVLTGLDGHGKIQOK 60
 QY 60 ABEAGITPOAYVDGMAGVKELMQLLDISYDKFIRTTDDYHEKVAQVFERLLAODDIY 119
 DB 61 IAEKXITPKAYVDEIVAGIKELMKNWNIYDKFIRTTDEYHVELVQKFKTYLDKGLYK 120
 QY 120 LGYSGWYSVDEEFFTESQLAEVFRDAGNVTGGIAPSGHEVWSESYFLRLSKYQDR 179


```

121 XSAEYGVCTPCBSFWTEQL-----VDGKCPDGRPVVEKTEKAEYFFKMSKYAD 170
180 RLVEFFKAHPEFTIPDGRLEMLNFIPEGLEDLAVSRITFTWGVVPVPSNPKHVYVWID 239
171 RLKYLIEDHDFIQPSRKNEMNLKPGQLDLCISRSFDMGIPITFDEKHYVYVWID 230
240 ALLNATYALGYAODEHGNFDPKFWNGTVFHMVGKDLIRPHSIYWPILLMMLDVKLDPRLTA 299
231 ALSNYITALGYGSDNDELXNKFPA DL-HLVGKDIIIRFTIYWPIMMALDLPKQVFG 289
300 HGFVMDKGVSKSGNVVYPEMLVERYGLDPLRYLYLMRNLPGVSDGTFPTEDYVGRINY 359
290 HGWLLVDGGMVSKSGNVVDPVVLNFEFGTDPVRYLLHEIPFGSDGFNNEIFIKKINS 349
360 ELANDLGNLNRVTSMINKYFDQIPAYVEGVTEFHDVLAEVAEKSIADPHTHMEAVDYP 419
350 DLANDLGNLVRRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLIEPFMMET 408
420 RALEAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLIEPFMMET 479
409 EALSHINDLKRANKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLIEPFMMET 468
480 SRAVLTOIGLEVSLSLENLADFPADVTVVAKGTPIPPLNMBEIEIAYKEQMEGNKPA 539
469 GEKIKTQINT-ELDTWESLSAFDGTAGTKVSKGEVIFPRIDVDKKEIEELNKLKEEQKA 527
540 VERKWNPDVELKLNKDEIKFEDPKVEISVAEVEKSVKVEGSKLLOFRLDAGDGEDRO 599
528 TRK-----MQLKPEISIDVDKDLRLRVKLECEPVKSKKLLKLVKELG-GEERQ 578
600 ILSGIAYYNEQELVGGKQVIVANLXPRKMKKYVSGQML--SABHDGKLTLLTVDP 657
579 VLSGISQFYKPE-DLICKVVLVANLAPAKLMGQ-LSQGMILAVATDDSDKLTLDIPED 636
658 VPNGSVI 664
637 IPTGSIV 643

```

RESULT 15

```

PCT-US02-09107B-53102
; Sequence 53102, Application pc/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53102
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Clostridium difficile
PCT-US02-09107B-53102

```

```

Query Match 45.2%; Score 1572; DB 1; Length 645;
Best Local Similarity 48.0%; Pred. No. 1.4e-77;
Matches 319; Conservative 113; Mismatches 205; Indels 28; Gaps 11;

1 MSEKNFYITPTIYPSGKLHIGSAITACDVARYKRLMGVDFYFLLTGLDEHGKIQOK 60
1 MSKPSFYVITPTIYPSGGLHIGHTYSTVAADTIAEFKFCGVYDKVFLAGTDEHGKIQK 60

```

Search completed: December 17, 2004, 22:12:06
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2004, 21:41:55 ; Search time 158 Seconds
(without alignments)
1509.842 Million cell updates/sec

Title: US-10-606-607-2

Perfect score: 3479

Sequence: 1 MSKKNFYITPTIYPSGKLH.....DGKLLILLTVDPAPNGSVIG 665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3479	100.0	665	2	AAW27660 Streptoco
2	3446	99.1	665	6	ABU01176 S. pneumo
3	3446	99.1	679	4	AAU37825 Streptoco
4	3446	99.1	679	6	ABU45969 Protein e
5	3429	98.6	679	8	ADK46236 Streptoco
6	2883.5	82.9	666	5	ABP26526 Streptoco
7	2883.5	82.9	666	5	ABU46491 Protein e
8	2883	82.9	669	6	ABU44499 Protein e
9	2732.5	78.5	665	5	ABP30434 Streptoco
10	2732.5	78.5	682	5	ABP26525 Streptoco
11	2392.5	68.8	662	5	ABBS4105 Streptoco
12	2295	66.0	669	6	ABU29885 Protein e
13	2231.5	65.9	674	7	ADC96574 E. faeciu
14	2288	65.8	677	4	AAU35070 Enterococ
15	2288	65.8	677	6	ABU44551 Protein e
16	2284.5	65.7	666	4	AAU33437 Enterococ
17	2154	61.9	680	6	ADB08140 Allostoc
18	2113	60.7	664	5	ABBS0126 Listeria
19	2113	60.7	664	6	ABU32461 Protein e
20	2040	58.6	652	6	ADB08138 Allostoc
21	1949	56.0	659	6	ABU43600 Protein e
22	1939	55.7	657	6	ABRS9494 S. aureus
23	1935.5	55.6	669	6	ABU18199 Protein e
24	1934	55.6	657	4	AAU36542 Staphyloc
25	1931	55.5	657	2	AAW20051 Methionyl

ALIGNMENTS

RESULT 1

AAW27660

ID AAW27660 standard; protein; 665 AA.

AC AAW27660;

DT 28-APR-1998 (first entry)

DE Streptococcus pneumoniae methionyl tRNA synthetase.

KW Methionyl tRNA synthetase; metS polypeptide; antibacterial; vaccine;
KW immune response; meningitis; bacterial infection.

OS Streptococcus pneumoniae.

PN WO9739012-A1.

PD 23-OCT-1997.

PF 18-APR-1997; 97WO-US006555.

PR 18-APR-1996; 96GB-00007999.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PA (SMIK) SMITHKLINE BEECHAM PLC.

PI Lawlor EJ;

DR WPI; 1997-526388/48.

DR N-PSDB; AAT68098.

XX Polynucleotide encoding Streptococcus pneumoniae methionyl tRNA
XX synthetase - useful to diagnose, treat or prevent bacterial infection,
XX especially meningitis.

PS Claim 12; Page 36-38; 47pp; English.

CC The present sequence represents methionyl tRNA synthetase (mets) from
CC Streptococcus pneumoniae. The polynucleotide encoding mets can be used
CC for the recombinant production of mets, which can be used to treat
CC conditions requiring mets activity. Sequences antisense to the
CC polynucleotide can be used to control mets expression. Mets, or mets
CC polynucleotide can be used to induce an immune response, i.e. an
CC antibody (Ab) and/or T cell response, against S. pneumoniae to protect
CC against infection, or to screen for antagonists or agonists of the
CC polynucleotide encoding mets or mets's activity, i.e. antibacterials. The
CC antagonist, e.g. an anti-mets Ab, can be used to treat conditions requiring
CC mets inhibition, e.g. an S. pneumoniae infection, particularly

ABU16009 Protein e
AAU37420 Staphyloc
ABR59493 S. aureus
ABM71940 Staphyloc
ABU43118 Protein e
ABP38505 Staphyloc
AAU34400 Staphyloc
ADB08136 Allostoc
AAG81883 S. epider
ABU24718 Protein e
ABU23723 Protein e
ABU25178 Protein e
ADH86149 Enterococ
ABU18430 Protein e
ABU26515 Protein e
ABU48823 Protein e
AAU39861 Helicobac
AAW20999 H. pylori
ABU30785 Protein e
AAW75910 Helicobac

26 1931 55.5 657 6 ABU16009
27 1930 55.5 657 4 AAU37420
28 1930 55.5 657 6 ABR59493
29 1930 55.5 657 6 ABM71940
30 1922.5 55.3 656 6 ABU43118
31 1922.5 55.3 660 5 ABP38505
32 1884 54.2 652 4 AAU34400
33 1856 53.3 605 6 ADB08136
34 1787.5 50.8 613 4 AAG81883
35 1678 48.2 674 6 ABU24718
36 1640.5 47.2 644 6 ABU23723
37 1572 45.2 645 6 ABU25178
38 1492.5 42.9 493 7 ADH86149
39 1383 39.8 670 6 ABU18430
40 1189.5 34.2 628 6 ABU26515
41 1153.5 33.2 521 6 ABU48823
42 1152.5 33.1 656 4 AAU39861
43 1152.5 33.1 662 2 AAW20999
44 1148.5 33.0 650 6 ABU30785
45 1137.5 32.7 648 2 AAW75910

CC meningitis. A fragment of the polynucleotide encoding mets can used as a
 CC probe to isolate full length or related sequences, or diagnose, e.g. by
 CC polymerase chain reaction, infection stage and type, including mutation
 CC and polymorphism detection. Diagnosis may also be achieved by detecting
 CC mets gene overexpression, e.g. by immunassay. The Ab can used to treat
 CC infection, isolate or identify mets expressing clones, purify mets and as
 CC an immunoassay reagent. More generally, the products can prevent adhesion
 CC of bacteria to wounds and in dwelling devices, block mets protein
 CC mediated invasion of mammalian cells and block the normal progression of
 CC infection
 XX
 SQ

Sequence 665 AA;

Query Match 100.0%; Score 3479; DB 2; Length 665;
 Best Local Similarity 100.0%; Pred. No. 2e-284;
 Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLTGLDEHGQKIQQK 60
 DB 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLTGLDEHGQKIQQK 60

QY 61 ABEAGTTPCAYVDGKAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAODDIYL 120
 DB 61 ABEAGTTPCAYVDGKAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAODDIYL 120

QY 121 GEYSGWYSDEEFFTESQLAEFVFRDEAGNVTGGIAPSGHEVWEVSSEYFLRLSKYQDR 180
 DB 121 GEYSGWYSDEEFFTESQLAEFVFRDEAGNVTGGIAPSGHEVWEVSSEYFLRLSKYQDR 180

QY 181 LVEFFKAHPEFTTPGRLNEMLENTIEPGLDLAVSRFTFTGVPVSPNKHVVVWIDA 240
 DB 181 LVEFFKAHPEFTTPGRLNEMLENTIEPGLDLAVSRFTFTGVPVSPNKHVVVWIDA 240

QY 241 LNNYATALGYAODEGNFQKFWNGTVFHMVGKDIILRFHSIYWPILLMMLDKVLPDLIAH 300
 DB 241 LNNYATALGYAODEGNFQKFWNGTVFHMVGKDIILRFHSIYWPILLMMLDKVLPDLIAH 300

QY 301 GWFVKDGMKSKGNVVPPEMLVRYGLDPLRYLYLMNLPVSGDGTTPEDYVGRINYE 360
 DB 301 GWFVKDGMKSKGNVVPPEMLVRYGLDPLRYLYLMNLPVSGDGTTPEDYVGRINYE 360

QY 361 LANDLGNLLNRVSMINKYDFDQIPAYVEGVTEPDHVLAEVAEKSIADPHTHMEADVPR 420
 DB 361 LANDLGNLLNRVSMINKYDFDQIPAYVEGVTEPDHVLAEVAEKSIADPHTHMEADVPR 420

QY 421 ALEAVWTLISRTNKYIDETAPVWLDKDEALRDQLASVMSHWQASIRVVVAHLIEPFFMETS 480
 DB 421 ALEAVWTLISRTNKYIDETAPVWLDKDEALRDQLASVMSHWQASIRVVVAHLIEPFFMETS 480

QY 481 RAVLTQIGLEEVSSLENISLADFPADVTVAAGTPIFRLNNEEBEIAIKQMEGNKPAV 540
 DB 481 RAVLTQIGLEEVSSLENISLADFPADVTVAAGTPIFRLNNEEBEIAIKQMEGNKPAV 540

QY 541 EKEMNPDEVELKLNKDEIKFDFKVEIRVAEVEKSVESGSKLLQRLDAGDGEDRQI 600
 DB 541 EKEMNPDEVELKLNKDEIKFDFKVEIRVAEVEKSVESGSKLLQRLDAGDGEDRQI 600

QY 601 LSGIAKYPNEBELGKVKQIVIANLKPRMKKKYVSQGMILSAEHGDKLTLLTVDPVFN 660
 DB 601 LSGIAKYPNEBELGKVKQIVIANLKPRMKKKYVSQGMILSAEHGDKLTLLTVDPVFN 660

QY 661 GSVIG 665
 DB 661 GSVIG 665

RESULT 2

ABU01176

ID ABU01176 standard; protein; 665 AA.

XX

AC ABU01176;

XX

DT 23-OCT-2003 (revised)

DT 11-FEB-2003 (first entry)
 XX S. pneumoniae type 4 strain protein from coding region #749.
 DE Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
 XX antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
 KW gene therapy; vaccine.
 XX Streptococcus pneumoniae; type 4 strain.
 OS
 XX WO200277021-A2.
 PN 03-OCT-2002.
 XX 27-MAR-2002; 2002WO-IB002163.
 PF 27-MAR-2001; 2001GB-00007658.
 PR (CHIR-) CHIRON SPA.
 XX (GENO-) INST GENOMIC RES.
 PA Masignani V, Tettelin H, Frazer C;
 XX WPI; 2003-040579/03.
 DR N-PSDB; ABX06461.
 XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
 XX ear infection.
 PS Claim 1; SEQ ID NO 1498; 56pp; English.
 XX
 CC The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC ABS5454. Also included are an antibody which binds one of the proteins,
 CC treating a patient by administering the protein, DNA or antibody (in a
 CC composition), a kit comprising first and second primers, which are the
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a
 CC sequence not defined in the specification, for amplifying a target
 CC sequence contained within a Streptococcus nucleic acid sequence, where
 CC the first primer is substantially complementary to the target sequence
 CC and the second primer is substantially complementary to the complement of
 CC the target sequence, and where the parts of the primers having
 CC substantial complementarity define the termini of the target sequence to
 CC be amplified, assay comprising contacting a test compound with the
 CC protein, and determining whether the test compound binds to the protein
 CC and a Streptococcus pneumoniae bacterium, where one or more genes
 CC encoding the proteins has been rendered inactive. The proteins, nucleic
 CC acid molecules, antibody and compositions are useful as medicaments for
 CC treating or preventing a disease or infection due to streptococcus
 CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
 CC media or ear infection. They are also useful in developing vaccines,
 CC diagnostics and antibiotics. The methods are useful for identifying
 CC immunodominant proteins. The present sequence is one of the 2469 proteins
 CC expressed by the identified coding regions from the genomic sequence.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 665 AA;

Query Match

99.1%; Score 3446; DB 6; Length 665;

Best Local Similarity 99.2%; Pred. No. 1.2e-281;

Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY

1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLTGLDEHGQKIQQK 60

DB

1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLTGLDEHGQKIQQK 60

Db 495 RAVLTQLGLEEVSSLENLSADFPADVTVAAGTPIFFRLNMEESIAVIKEQMEGNKPAV 554
 Qy 541 EKEWNPDEVELKLNKDEIKFDPDKVEIRVAVKESVKEGSDKLLQRLDAGDGEDROI 600
 Db 555 EKEWNPDEVELKLNKDEIKFDPDKVEIRVAVKESVKEGSDKLLQRLDAGDGEDROI 614
 Qy 601 LSGIAKYYPNEQELGKVKQIVANLKPROMKKYVSQGMILSAEHDGKLLTLTVDPAVNP 660
 Db 615 LSGIAKYYPNEQELGKVKQIVANLKPROMKKYVSQGMILSAEHDGKLLTLTVDPAVNP 674
 Qy 661 GSVIG 665
 Db 675 GSVIG 679

RESULT 4
 ABU45969 standard; protein; 679 AA.
 AC ABU45969;
 XX 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #31496.
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Streptococcus pneumoniae.
 OS Wo20027183-A2.
 PN 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 23-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Irawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA49839.

Claim 25; SEQ ID NO 73893; 1766pp; English.
 The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 679 AA;

Query Match 99.1%; Score 3446; DB 6; Length 679;
 Best Local Similarity 99.2%; Pred. No. 1.3e-281;
 Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGYDVLYLTGLDEHGKIQK 60
 Db 15 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGYDVLYLTGLDEHGKIQK 74
 Qy 61 ABEAGITPQAYVDGMAGVVKELWQLLDISYDKFIRTTDDYHEKWAQVFERLLAODIYL 120
 Db 75 ABEAGITPQAYVDGMAGVVKELWQLLDISYDKFIRTTDDYHEKWAQVFERLLAODIYL 134
 Qy 121 GEYSGWYSVSDDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDR 180
 Db 135 GEYSGWYSVSDDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDR 194
 Qy 181 LVEFFKAHPEFITPDGRLEMLNFIPEGLEDLAVSRITFTGWVPVSPKHYVYVWIDA 240
 Db 195 LVEFFKAHPEFITPDGRLEMLNFIPEGLEDLAVSRITFTGWVPVSPKHYVYVWIDA 254
 Qy 241 LLNYATALGYAODEHGNFDFKFNVTGVHVGKDIILRFHSIYWPILLMLDLVKLPORLIAH 300
 Db 255 LLNYATALGYAODEHGNFDFKFNVTGVHVGKDIILRFHSIYWPILLMLDLVKLPORLIAH 314
 Qy 301 GWFVMKDGKMSKSGNVVYPEMLVERISGLDPLRYILMRNLVPGSGDTFTPDYVGRINYE 360
 Db 315 GWFVMKDGKMSKSGNVVYPEMLVERISGLDPLRYILMRNLVPGSGDTFTPDYVGRINYE 374
 Qy 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSTADPHTHMEAVDYPR 420
 Db 375 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSTADPHTHMEAVDYPR 434
 Qy 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDQASVMSHWQASIRVAHLIEPFMMETS 480
 Db 435 ALBAVWTLISRTNKYIDETAPWLDKDEALRDQASVMSHWQASIRVAHLIEPFMMETS 494
 Qy 481 RAVLTQIGLEEVSSLENLSADFPADVTVAAGTPIFFRLNMEESIAVIKEQMEGNKPAV 540
 Db 495 RAVLTQIGLEEVSSLENLSADFPADVTVAAGTPIFFRLNMEESIAVIKEQMEGNKPAV 554
 Qy 541 EKEWNPDEVELKLNKDEIKFDPDKVEIRVAVKESVKEGSDKLLQRLDAGDGEDROI 600
 Db 555 EKEWNPDEVELKLNKDEIKFDPDKVEIRVAVKESVKEGSDKLLQRLDAGDGEDROI 614
 Qy 601 LSGIAKYYPNEQELGKVKQIVANLKPROMKKYVSQGMILSAEHDGKLLTLTVDPAVNP 660
 Db 615 LSGIAKYYPNEQELGKVKQIVANLKPROMKKYVSQGMILSAEHDGKLLTLTVDPAVNP 674
 Qy 661 GSVIG 665
 Db 675 GSVIG 679

RESULT 5
 ADK46236
 ID ADK46236 standard; protein; 679 AA.

XX ADK46236;
XX 20-MAY-2004 (first entry)
XX Streptococcus pneumoniae protein, Seq ID No 2751.
DE Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
KW Streptococcus pneumoniae.
OS Streptococcus pneumoniae.
XX US6699703-B1.
XX 02-MAR-2004.
XX 26-MAY-2000; 2000US-00583110.
XX 02-JUL-1997; 97US-0051553P.
PR 12-MAY-1998; 98US-0085131P.
PR 30-JUN-1998; 98US-00107433.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
PI WPI; 2004-212399/20.
XX N-PSDB; ADK43575.
XX New nucleic acid molecules and polypeptides useful for diagnosing,
PT preventing and treating pathological conditions resulting from bacterial
PT infection, e.g. Streptococcus pneumoniae infection, and in drug
PT screening.
XX
XX Disclosure; SEQ ID NO 2751; 301pp; English.
XX
XX The invention relates to isolated Streptococcus pneumoniae nucleic acids
CC and polypeptides. The nucleic acids and proteins are useful for
CC diagnosing, preventing and treating pathological conditions resulting
CC from bacterial infection, such as S. pneumoniae infection. These may also
CC be used for drug screening procedures. The present sequence represents a
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
CC data for this patent did not appear in the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
XX Sequence 679 AA;
Query Match 98.6%; Score 3429; DB 8; Length 679;
Best Local Similarity 98.8%; Pred. No. 3.4e-280;
Matches 657; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 1 MSEKNFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVFYLTGLDEHGQKIQQK 60
DB 15 MSEKNFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVFYLTGLDEHGQKIQQK 74
QY 61 ABEAGITPOAYVDGMAGVYKELQDLIDISYDKETRTDDYHEKVAQVFERLLAQDDIYL 120
DB 75 ABEAGITPOAYVDGMAGVYKELQDLIDISYDKETRTDDYHEKVAQVFERLLAQDDIYL 134
QY 121 GEYSGWYSVSDDEFFFTESQALAEFRBAGNVTGGIAPSGHEVEWVSEESYFLRLSKYQDR 180
DB 135 GEYSGWYSVSDDEFFFTESQALAEFRBAGNVTGGIAPSGHEVEWVSEESYFLRLSKYQDR 194
QY 181 LVFEFFKAHPFITPDGRLNEMLRNFTPEGLDIAVSRITPTWGPVPSNPKHVVYWIIDA 240
DB 195 LVFEFFKAHPFITPDGRLNEMLRNFTPEGLDIAVSRITPTWGPVPSNPKHVVYWIIDA 254
QY 241 LLNVATALGAQDEHGNFDFWNGTVPFHWGKIDILRPHSIYWPILLMLMDLKLDPDLIIAH 300
DB 255 LLNVATALGAQDEHGNFDFWNGTVPFHWGKIDILRPHSIYWPILLMLMDLKLDPDLIIAH 314
QY 301 GWFVKDGMKSKGNVYVPEMLVERYGLDPLRYLMLNLPVGSDDGTFTPEDYVGRINYE 360

DB 315 GWFVKDGMKSKGNVYVPEMLVERYGLDPLRYLMLNLPVGSDDGTFTPEDYVGRINYE 374
QY 361 LANDGLNLRARTVSMINKYFDGQIPAYVEGVTFEDHVLAEVAEKSIADFEHMEAVDYP 420
DB 375 LANDFGNLRARTVSMINKYFDGQIPAYVEGVTFEDHVLAEVAEKSIADFEHMEAVDYP 434
QY 421 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLIEFPMMETS 480
DB 435 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHLAASIRVVAHLIEFPMMETS 494
QY 481 RAVITQIGLEBVSSLENLSLADPPADVTVAAGTPIPPRLNMBEEIAYIKEQMEGNKPAV 540
DB 495 RAVITQIGLEBVSSLENLSLADPPADVTVAAGTPIPPRLNMBEEIAYIKEQMEGNKPAV 554
QY 541 EKEWNPDEVELKLNKDEIKFEDFKVEIRVAEVEKSVKSGDKLQFRLDAGDGEDRQI 600
DB 555 EKEWNPDEVELKLNKDEIKFEDFKVEIRVAEVEKSVKSGDKLQFRLDAGDGEDRQI 614
QY 601 LSGIAKYYPNEQELVGVKQIVANLKPRKMMKKYVSQGMILSAEHDGKLTLLTVDPVAPN 660
DB 615 LSGIAKYYPNEQELVGVKQIVANLKPRKMMKKYVSQGMILSAEHDGKLTLLTVDPVAPN 674
QY 661 GSVIG 665
DB 675 GSVIG 679
RESULT 6
ABP26526
ID ABP26526 standard; protein; 666 AA.
XX
AC ABP26526;
XX
DT 02-JUL-2002 (first entry)
XX Streptococcus polypeptide SEQ ID NO 2228.
DE Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX Streptococcus pyogenes.
OS
XX WO200234771-A2.
XX
XX 02-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-GB004789.
XX
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
PI Tettelin H;
XX
XX WPI; 2002-352536/38.
DR N-PSDB; ABN67157.
XX
PT New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
PS Claim 1; Page 3374; 4525pp; English.
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus (GBS (Streptococcus agalactiae) or group A streptococcus (GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and

CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX
 SQ Sequence 666 AA;

Query Match 82.9%; Score 2883.5; DB 5; Length 666;
 Best Local Similarity 81.9%; Pred. No. 3.9e-234;
 Matches 546; Conservative 54; Mismatches 64; Indels 3; Gaps 2;

QY 1 MSEKPYITPTIYPSGKLHIGSAVTTTACDVLARYKELMGVDYFYLTLDEHGQKIQOK 60
 Db 1 MKKKPFYITPTIYPSGKLHIGSAVTTTACDVLARYKELMGHEVYFYLTLDEHGQKIQTK 60

QY 61 ABEAGITPQAYVDGMAGVKELWQLLDISYDKFIRTTDDYHEKVAQVAFERLLAQDDIYL 120
 Db 61 AKEAGITPQYVDNMAKVAKWQLLDISYDFIRTTDDYHEVVAVFEKLLAQDDIYL 120

QY 121 GEYSGWYSVDEBFTESQLAVFDEAGNVGTGGTAPSGHEVWVSESYFLRLSKYQDR 180
 Db 121 GEYSGWYSVDEBFTESQLKEVFRDEQGVGTGGTAPSGHEVWVSESYFLRLSKYQDR 180

QY 181 LVEFFKAHPEFTTPQRLNEMLRNIEPGLDLAVSRITFTWGVVPVNPKHVVVWIDA 240
 Db 181 LVAFKERPDFTQPGRNEMVKNIEPGLDLAVSRITFTWGVVPVNPKHVVVWIDA 240

QY 241 LNYATALGYAQDEHGNDFKNGTVFHMVGKDIIRPHSIYWPILLMLMDVKLPDLRIAH 300
 Db 241 LNYATALGYRQANHANFDFKNGTVFHMVGKDIIRPHSIYWPILLMLMDLPMPDLRIAH 300

QY 301 GFVWKGDKGWSKGNVYVPEMLVRYGLDPLRYVLMENLPGVSGTFTPDYVGRINYE 360
 Db 301 GFVWKGDKGWSKGNVYVPEMLVRYGLDPLRYVLMENLPGVSGTFTPDYVGRINYE 360

QY 361 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADFHCHMEADVYPR 420
 Db 361 LANDLGNLLNRTVAMINKYFDGQIPAYVDNGTAFADLSQLIDAQADYVHKHMEADVYPR 420

QY 421 ALEAVWTLISRTNKYIDETAPNVLKDEALRDOLASVMSHWQASIRVVAHLIEPFFMETS 480
 Db 421 ALEAVWTLIARTNKYIDETAPNVLAKEDGDKAQOLASVMAHLAASRLVAHVQIPFFMETS 480

QY 481 RAVLTQIGLEEVSSLENISLADFPADVTVAKGTPIFPRLNWEEBIAIYKEQMEGNKPAV 540
 Db 481 AAIMQGLGLEPVSLSLALADFPANTKVAKGTPIFPRLDMEABIDYIKQM-GDSSAI 539

QY 541 --EKWNPDEVELKUNKBIKEDFDKVEIRVAEVKESVKEGSKLLQFLRLDAGDGEDR 598
 Db 540 SOKEWNPDEVALKSEKDVITETFDVAEIRVAEVKESVKEGSEKLLRFRVADAGDGQDR 599

QY 599 QILSGIAKYPNQEQLVGKGVQIVANLPRKMKKKYVSGQMLSAEHDKGLTLTVDPV 658
 Db 600 QILSGIAKYPNQEQLVGKGLQIVANLPRKMKKKYISQGMILSAEHGQDLTVLTVDSSV 659

QY 659 PNGSVIG 665
 Db 660 PNGSIIG 666

RESULT 7

ABU46491

ID ABU46491 standard; protein; 666 AA.

XX

AC ABU46491;

XX 19-JUN-2003 (first entry)
 DT Protein encoded by Prokaryotic essential gene #32018.
 XX
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Streptococcus pyogenes.
 KW
 XX WO200277183-A2.
 XX 03-OCT-2002.
 PD
 XX 21-MAR-2002; 2002WO-US009107.
 XX
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Chlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-0299926/02.
 DR N-PSDB; ACA50361.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 74415; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than S. aureus, S. typhimurium,
 CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: the sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 666 AA;

Query Match 82.9%; Score 2883.5; DB 5; Length 666;

Best Local Similarity 81.9%; Pred. No. 3.9e-234;

Matches 546; Conservative 54; Mismatches 64; Indels 3; Gaps 2;

Qy	1	NSBKPNFYITPTIYPSGKLHI	GSAYTTIACDVLARYKRLMGYDVFYITGLDEHGKIQOK	60	
Db	1	MMKSPFYITPTIYPSGKLHI	GSAYTTIACDVLARYKRLMGHEVYITGLDEHGQIKTK	60	
Qy	61	ASEAGITPOAYVDGMAVGKSLW	LLDISYDKFIRTTDDYHEKVQAVQPERILAAQDDIYL	120	
Db	61	AXEAGITQTYVDNMAKDVKAL	QQLLDISYDTFIRTTDDYHEEVAAPFEKLLAQDDIYL	120	
Qy	121	GSYSGWYSVDSBEFTTESQLAE	VFRDEAGNVTTGGIAPSGHEVEKVSSESYFLSLKYQDR	180	
Db	121	GSYSGWYSVDSBEFTTESQLKE	VFRDEDDQVIGGIAPSGHEVEKVSSESYFLSLKYDDR	180	
Qy	181	LYVEFFKAPERTTPOGRINEMLR	FIPEGLIEDLAVSRITTTTGWGPVPSNPKHVYVWIDA	240	
Db	181	LVAFKPERPDTQPGRNEMWKFI	PEGLIEDLAVSRITTTTGWGPVPSDPGHVYVWIDA	240	
Qy	241	LLNYATALGYAQDEHGNDFKWN	GVTFVHMVGKDIILRFHSIYWPIILLMMLDKVLPDRLIAH	300	
Db	241	LLNYATALGYROANHANFDKFN	GVTFVHMVGKDIILRFHSIYWPIILLMMLDLPMPDRLIAH	300	
Qy	301	GFWMKDGKMSKGNVY	PEMLVERGLDPLRYIYLAENLPVSGDGTFTPEVDVGRINYE	360	
Db	301	GFWMKDGKMSKGNVY	PEMLVERGLDPLRYIYLAENLPVSGDGTFTPEVDVGRINYE	360	
Qy	361	LANDLGNLNRVTSMINKY	FDGQIPAYVEGVTBFDHVLAEVASKSIADFTHTHMEAVDYPR	420	
Db	361	LANDLGNLNRVTSMINKY	FDGTPVAYVDNCTAFDADLSQIDAQLADYHKHMEAVDYPR	420	
Qy	421	ALEAVNTLI	SRNKYIIDETA	PWJLDKDBALBDQLASVMSHWQASIRVVAHLI	480
Db	421	ALEAVNTLI	ARTNKYIIDETA	PWJLAKEDGKAQTLASVVAHLASRLVAHVIOFPFMMETS	480
Qy	481	RAVLTIQIGLEEYSSLENISL	ADFADVTVAKGTPIPPRLNMSEETAYIKEQMEGNKPAV	540	
Db	481	AAIMAQLGLEPVSDDLSTL	ALADFPANTKVAKGTPIPPRLDMEAEIDYIKAQM-GDSSAI	539	
Qy	541	--EKWNVDEVELKUNKDEI	KPEDFDKVETIRVAEKEVSKVSGSKLLQPRLDAGCGEGR	598	
Db	540	SOQEKWVPEEVALKEKQVI	TFETFDVAEIRVAEKEVSKVSGSEKLLRFRVDAAGDQDR	599	
Qy	599	QILSGIAXYPNQEQLVGKKVQI	VANLKPROMKKYYSQGMILSAEHGDKLTLLTVDPAV	658	
Db	600	QILSGIANFYPNQEQLVGKKI	QIIVANLKPROMKKYYSQGMILSAEHGDLTVLTVDSSV	659	
Qy	659	PNGSVIG	665		
Db	660	PNGSIIG	666		

RESULT 8

ABU44499
ID ABU44499 standard; protein: 669 AA.

AA
AC ABU44499:

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #30026.

Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Streptococcus mutans.

AA
PN
WO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001: 2001US-00815242.

08-SEP-2001; 2001US-00948553.
25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072831.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L., Zamudio C., Malone C., Haselbeck R., Ohlson KL., Zyskind JW:

Wally L, Zamudio C, Malone C, Haseideck K, Christensen NJ, Zyskind
Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI: 2003-029926/02.

N-PSDB; ACA48369;

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 25: SEO ID NO 72423: 1766pp: English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:

- (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) proliferating a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at <http://wipo.int/pub/published/pct/sequences>

Sequence 669 AA;

Query Match	82.9%	Score 2883	DB: 6	Length 669
-------------	-------	------------	-------	------------

Query Match 02.3%, Score 2003, 22.07,
Best Local Similarity 82.4%; Pred. No. 4.3e-234;

Seq. Local Similarity	Seq. No.	Indels	Gaps
Matches 553; Conservative 46; Mismatches 64;	231	Indels 8;	Gaps 5;

1 MSEKN-FYITTPYYPSGKLHIGSAYTTIACDVLARYKRLMGYDVFYLTGLDEHGQKIQQ 59

1 MTEKQPFYITTPYYPSGKLHIGSAYTTIACDVLARYKRMNHDVFYLTGLDEHGQKIQQ 60

60 KAEFAGITPOAYVDGMAVGKELWOLLDISYDKFIRTTDDYHEKVVAQVFERLLAQDDIY 119

61 KSEEGITPOAYVDGMAVGVKELWKLDSYDKFIRTTDDYHEKVVADVFEKLLAOGDIY 120

120 IGEYSGWYSVSDEEFFETESOI.AEVRDEAGNVTGGIAPSGHEVWVSEESYFLBLISKYOD 179

121 LGEYSGWYSVSDDEFFETESOLEEIPRDEDEGKVIIGGIAPSGCHEVEWVSEESYFLRLSNYAD 180

180 RI.VEEFKAHPEFTTPDGRINEMIRNFTERPGLIEDI.AVSRTTFTWGVPPSPNPKHVVVYVWD 239

181 PLYDEEHAHPDEIOPDGBNNEIKNETEREGLEDIAVSPTSETWGVKVPSPKHWVWTD 240

240 AII NVATAI GYA ODEH GNE DK EWN CTVEH MVG KDT I REH STYWP T I MML DVKI PDRI TA 299

341 31 INVA TA1 CYCOPRA HANER KENCTVEHMYCKDII PEHSI YWBII MMID KXDEPIA 300

CC 300 HGWFMKDKGMSKSGNVVYPMELVRYGLDPLRYLMLNLPVSGDGTFTPEDYVGRINY 359
 CC 301 HGWFMKDKGMSKSGNVVYPMELVRYGLDPLRYLMLNLPVSGDGTFTPEDYVGRINY 360
 CC 360 ELANDGLNLLNRTVSMINKYFGQIPAYVGVTEBDHVAEVAEKSIAADFTHEANDYD 419
 CC 361 ELANDGLNLLNRTVSMINKYFGQIPAYVGVTEBDHVAEVAEKSIAADFTHEANDYD 420
 CC 420 RALEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVNSHWQASIRVVAHLIPEPFMMET 479
 CC 421 RALEVWNLISRTNKYIDETAPWLDKDEALRDQLASVNSHWQASIRVVAHLIPEPFMMET 480
 CC 480 SRAVLTOIGLEBVSSLENIUSLADFPADVTVVAKGTPPIPRLNMBEIEIAYIKQWEGNKPA 539
 CC 481 SNAIMEQLGLAGQFLEGLLGLAGLPTGVKVIAGKTPPIPRLDMDIEIAYIKQW-GNSA 539
 CC 540 V----EKWNPDEVELKLNKDIKPEDPKVIRVAEVEKSVKSGDKLLQFLRDAGDG 595
 CC 540 IQAEEKEWNPVTVLKNKDIKPEDPKVIRVAEVEKSVKSGDKLLQFLRDAGDG 599
 CC 596 EDRQILSGIAKYYPNEQLVGGKQIVANLKPRLKMMKKYVSGMILSAEH-DGKLTLLTV 654
 CC 600 EDRQILSGIAKYYPNEQLVGGKQIVANLKPRLKMMKKYVSGMILSAEHADGKLTLLTV 658
 CC 655 DPAVNGSVIG 665
 CC 659 DSSVNGSQIG 669

RESULT 9
 ID ABP30434 standard; protein; 665 AA.
 XX ABP30434;
 XX 02-JUL-2002 (first entry)
 XX Streptococcus polypeptide SEQ ID NO 10044.
 XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 XX group A streptococcus; Streptococcus pyogenes; antibacterial;
 XX antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX Streptococcus agalactiae.
 XX WO200234771-A2.
 XX 02-MAY-2002.
 XX 29-OCT-2001; 2001WO-GB004789.
 XX 27-OCT-2000; 2000GB-00026333.
 XX 24-NOV-2000; 2000GB-00028727.
 XX 07-MAR-2001; 2001GB-00005640.
 XX (CHIR-) CHIRON SPA.
 XX (GENO-) INST GENOMIC RES.
 XX Telford J, Maignani V, Margarit Y Rosi, Grandi G, Fraser C;
 XX Tettelin H;
 XX WPI; 2002-352536/38.
 XX N-PSDB; ABN71065.
 XX New Streptococcus protein for the treatment or prevention of infection or
 XX disease caused by Streptococcus bacteria, such as meningitis, and for
 XX detecting a compound that binds to the protein.
 XX Claim 1; Page 4133-4134; 4525pp; English.
 XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 CC Sequence 665 AA;
 CC Query Match 78.5%; Score 2732.5; DB 5; Length 665;
 CC Best Local Similarity 78.2%; Pred. No. 2.2e-221;
 CC Matches 519; Conservative 61; Mismatches 81; Indels 3; Gaps 3;
 CC
 CC 3 EKNFYTTPIYPSGKHLHIGSAITTTACDVLARYKRLMGYDVFYLTGLDEHGGKIQKAE 62
 CC 4 KKSFYTTPIYPSGKHLHIGSAITTTACDVLARYKRLMGYDVFYLTGLDEHGGKIQKAE 63
 CC 63 EAGITPQAVYDGMVGVKELWLLDISYDKFIRTTDDYHEKVVAVQVFERLLAQDDIYLGE 122
 CC 64 EAGITPQAVYDGMVGVKELWLLDISYDKFIRTTDDYHEKVVAVQVFERLLAQDDIYLGE 123
 CC 123 YSGWYSVSDSEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRSKTDRLV 182
 CC 124 YTGWYSVSDSEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRSKTDRLV 183
 CC 183 EPPKAHPEFTPDGRNEMLRNFIPEGLDGLAVSRFTTGTWGPVSPNPKHVVWIDALL 242
 CC 184 AYYAEHPEFTPDGRNEMLRNFIPEGLDGLAVSRFTTGTWGPVSPNPKHVVWIDALL 243
 CC 243 NYATALGYA-QDEHGNFDFWNGTVFHMVGKIDILRPHSYIWPILLMMLDVKLPDLRIANG 301
 CC 244 NYISALGYWSDDLISQYHKEWPAI-HMIGKIDILRPHSYIWPILLMMLDVKLPDLRIANG 302
 CC 302 WFWMDGKMSKSGNVVYPMELVRYGLDPLRYLMLNLPVSGDGTFTPEDYVGRINYEL 361
 CC 303 WFWMDGKMSKSGNVVYPMELVRYGLDPLRYLMLNLPVSGDGTFTPEDYVGRINYEL 362
 CC 362 ANDLGNLLNRTVSMINKYFDGQIPAYVGVTEBDHVAEVAEKSIAADFTHEANDYDPA 421
 CC 363 ANDLGNLLNRTVSMINKYFDGQIPAYVGVTEBDHVAEVAEKSIAADFTHEANDYDPA 421
 CC 422 LEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVNSHWQASIRVVAHLIPEPFMMETSR 481
 CC 422 LEAVWNLISRTNKYIDETAPWLDKDEALRDQLASVNSHWQASIRVVAHLIPEPFMMETSD 481
 CC 482 AVLTOIGLEBVSSLENIUSLADFPADVTVVAKGTPPIPRLNMBEIEIAYIKQWEGNKPAVE 541
 CC 482 AIMEQLGLAGTFLDKLFTADLPFGVVRVAKGSPIFPRLDMDIEIAYIKQWEGNKPAVE 541
 CC 542 KENPDEVELKLNKDIKPEDPKVIRVAEVEKSVKSGDKLLQFLRDAGDGQRQL 601
 CC 542 KENPDEVELKLNKDIKPEDPKVIRVAEVEKSVKSGDKLLQFLRDAGDGQRQL 601
 CC 602 SGIAKYYPNEQLVGGKQIVANLKPRLKMMKKYVSGMILSAEHADGKLTLLTVDPVNG 661
 CC 602 SGIAKYYPNEQLVGGKQIVANLKPRLKMMKKYVSGMILSAEHADGKLTLLTVDPVNG 661
 CC 662 SVIG 665
 CC 662 SIIG 665
 CC
 CC RESULT 10
 CC ABP26525

ID ABP26525 standard; protein; 682 AA.
 XX AC ABP26525;
 XX DT 02-JUL-2002 (first entry)
 XX DE Streptococcus polypeptide SEQ ID NO 2226.
 XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX OS Streptococcus agalactiae.
 XX WO200234771-A2.
 XX PD 02-MAY-2002.
 XX PF 29-OCT-2001; 2001WO-GB004789.
 XX PR 27-OCT-2000; 2000GB-00026333.
 PR 24-NOV-2000; 2000GB-00028727.
 PR 07-MAR-2001; 2001GB-00005640.
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
 PI Tettelin H;
 XX WPI: 2002-352536/38.
 DR N-PSDB; ABN67156.
 XX New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.
 XX Claim 1; Page 3374; 4525pp; English.
 XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX Sequence 682 AA;
 SQ
 Query Match 78.5%; Score 2732.5; DB 5; Length 682;
 Best Local Similarity 78.2%; Pred. No. 2.3e-221;
 Matches 519; Conservative 61; Mismatches 81; Indels 3; Gaps 3;
 QY 3 EKNEYITPIYPSGKLHIGSAVTTIACDVLARVRLMGVYDVLTLGLDHEGKIOKAE 62
 DB 21 KKSFYITPIYPSGKLHIGSAVTTIACDVLARVRLMGVYDVLTLGLDHEGKIOKAE 80
 QY 63 EAGITPQAYVDGMAVGKELWQLLDISYDKFIRTTDDYHEKVAQVPERLLAQDDIYLGE 122
 DB 81 EAGITPQEQYDGMASVKTWELLDDISYDKFIRTTDYHEEAVAKIFEQLAQDDIYLGE 140
 QY 123 YSGWYSVDEBFFFTESQLAEVYRDENGNGMIGVAPSGHEVEKVSSESYFFRMSKYADELK 182

DB 141 YTGWYSVDEBFFFTESQLAEVYRDENGNGMIGVAPSGHEVEKVSSESYFFRMSKYADELK 200
 QY 183 EFFKAHPEFITPDGRLNEMLRNFTEPGLDLAVSRITFTWGPVPSNPKHVVYVWIDALL 242
 DB 201 AYAHEPEFIQPDGRNEMLRNFTEPGLDLAVSRITFTWGPVPSNPKHVVYVWIDALM 260
 QY 243 NYATALGYA-QDEHGNFDKFWNGTVFHMVGDIILRFHSIYWPILLMLMDVKLPDLRIAHG 301
 DB 261 NYISALGYCWSDDLSSQYHKFWPADI-EMIGKDIILRFHSIYWPIMLMALDLPKRLVAHG 319
 QY 302 WFWMDGKWSKSGNVVYPEMLVERYGLDPLRYILMRNLPGVSDGTFTPTEDYVGRINVEL 361
 DB 320 WFWMDGKWSKSGNVVYPEMLVERFGLDPLRYILMRSLPGVSDGTFTPTEDYVGRINVEL 379
 QY 362 ANDLGNLNRVTSMINKYFDQIQIPAYVEGVTFFDHVLAEVAEKSIADPHHMEAVDYFPA 421
 DB 380 ANDLGNLNRITAMVNYKFDGEVPRFAV-ATDFDADLASVATDSTENYHKQMEAVDFPRA 438
 QY 422 LEAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHWOASIRVVAHLIEPFNMTSR 481
 DB 439 LEAVWTLISRTNKYIDETAPWLVAKDETDRDKLAAMSHLVASLRVVAHLIQPFNMTSD 498
 QY 482 AVLTOIGLEEVSSLENLSLADFPADTVVAKGTPFPRLNMEEEIAYIKEQMEGNKPAVE 541
 DB 499 ALMEQGLGATFDELEKLTFAADLPEGVRVVAKSGPIFFPELDMEDEITYIKEQMNAGKAPVE 558
 QY 542 KEWNPDEVELKLNKDEIFEDFDKVEIRVAEVEKSVKVEGSKLLOPRLDAGDGEDRQIL 601
 DB 559 KEWNPDEVELTSSKGQIKFEDFDAYEIRVAEVEKVEGSKLRLRFLDAGDEGHRQIL 618
 QY 602 SGIAKYFNEQELVGKQVQIVANLKPROMKKYVSCGMILSAEHDGKLTLLTVDPAVNG 661
 DB 619 SGIAKYFNEQELVGKQVQIVANLKPROMKKYVSCGMILSAEHDGKLTLLTVDPAVNG 678
 QY 662 SVIG 665
 DB 679 SIIG 682
 RESULT 11
 ABBS54.05
 ID ABBS54105 standard; protein; 662 AA.
 XX AC ABBS54105;
 XX DT 29-AUG-2003 (revised)
 DT 16-MAY-2002 (first entry)
 DE Lactococcus lactis protein metS.
 KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 OS Lactococcus lactis; IL1403.
 XX FR2807446-A1.
 XX PD 12-OCT-2001.
 XX PF 11-APR-2000; 2000FR-00004630.
 XX PR 11-APR-2000; 2000FR-00004630.
 XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PA Bolotine A, Scroquine A, Renault P, Ehrlich SD;
 PI WPI; 2002-043418/06.
 DR New nucleotide sequence useful in the identification of Lactococcus
 PT lactis and related species.
 XX Claim 6; SEQ ID NO 807; 2504pp; French.
 XX

CC The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (AB53300-AB55621). The nucleic
 CC acid sequence is useful in the detection and/or amplification of nucleic
 CC acid sequence, particularly to identify Lactococcus lactis or related
 CC species. The proteins of the invention are useful for the biosynthesis or
 CC biodegradation of a composition of interest. The invention helps research
 CC in lactic bacteria, particularly useful in the production of yogurt and
 CC cheese. Note: The sequence data for this patent is based on equivalent
 CC patent WO200177334 (published 18-OCT-2001) which is available in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
 CC standardise OS field)
 CC
 CC SQ Sequence 662 AA;

Query Match 68.8%; Score 2392.5; DB 5; Length 662;
 Best Local Similarity 68.8%; Pred. NO. 1.1e-192;
 Matches 458; Conservative 88; Mismatches 113; Indels 9; Gaps 7;

QY 1 MSF-KNPIYITPTIYPSGKLIHSGSAYTTIACDVLARYKELMGYDVFYLTGLDGHGQKIQ 59
 DB 1 MTENKTFYITPTIYPSGKLIHSGSAYTTIACDVLARYKELMGYDVFYLTGLDGHGQKIQ 60

QY 60 KAREAGITPOAYVDGMAVGKELWQLLDISYDFIRTTDDYHEKVAQVFERLLAQDDIY 119
 DB 61 KARELGWTPKEYLDPVADVOELWKLDDISYDFIRTTDDYHEEAVAKAFELQLDQDIY 120

QY 120 LGSYGWSVSDDEFTESQLAELVDFDEAGNVTGGIAPSGCHEVWVSESYFLRSKYQD 179
 DB 121 LGYAGWSVSDDEFTETQLEIFRDESGNITGGIAPSGCHEVWVSESYFLRSGKYAD 180

QY 180 RLVEFKAPHEFTTPDGRNLEMLNRIEFGLEDLAVSRITFTWGVPSNPKHVVYVWD 239
 DB 181 WLQYDDEHPDFQPEVRKENVNINIKFGLDGLALTRTSFTWGVPSNPKHVVYVWFD 240

QY 240 ALLNYATAGYAQDEHGFNFKWNGVFMVGVKDLRHSIYWPILLMMLDKLPLDLIA 299
 DB 241 ALLNYATAGYNSDNDNFKKYWPG--INMGKEIVRFTHTIYWPIMLHALGPAKKIFA 298

QY 300 HGFVMDKQWSKSGKNVYVPELVRYGLDPLRYILMRNLVPGSDGTTPEDYVGRNY 359
 DB 299 HGWLLMKDQWSKSGKNVYVPELVRYGLDPLRYILMRNLVPGSDGTTPEDYVGRNY 358

QY 360 ELANDLGNLLNRVSMINKYFDGQIPAYVGV--TEPDHVLAEVAKSIADFTTHMEAVDY 418
 DB 359 DLANDLGNLLNRVSMINKYNDGKIEA--TGVSTEPDASLEEVVEETISHFKAMDKEF 416

QY 419 FRALEAVWTLISSTNKYIDETAPWLDKDBALRDQLASVNSHWQASIRVVAHLISPFME 478
 DB 417 NVALADWTLISSTNKYIDETAPWLDKDBALRDQLASVNSHWQASIRVVAHLISPFME 476

QY 479 TSRAVLTOIGLEVS--SLENLSLADPADVTVVAKTPIFPRLNMEETAYIKMEQGNK 537
 DB 477 TSKGITEQLGNDERSFLENLSG--YSFTHPVVAKGQPIPRLDVEEEVAYIKLQVAGG- 534

QY 536 PAYEKEWNPDEVELKLNKDBIKFDPEDKVEIRVAEVKEYSKVSGDKLQPLDAGDGD 597
 DB 535 VLPEKEWNPDEVELNLTLPQIKFDDFEKIELKVAEVLVEPEVSGDKLRLFKLDAGDSEP 594

QY 598 RQLSGIAKYYPNEQLVGVKQIVANLPRKQMKKYVSGMLSNRHDCKLTILTVDDPA 657
 DB 595 RQLSGIAKYYPNEQLVGVKQIVANLPRKQMKKYVSGMLSNRHDCKLTILTVDDDD 654

QY 658 VPNGSVIG 665
 DB 655 VPAGSLIG 662

RESULT 12
 ABU29885
 ID ABU29885 standard; protein; 669 AA.
 XX
 AC ABU29885;

XX DT 19-JUN-2003 (first entry)
 XX DE Protein encoded by prokaryotic essential gene #15412.
 XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX OS Enterococcus faecium.
 XX FN WO200277183-A2.
 XX PD 03-OCT-2002.
 XX PF 21-MAR-2002; 2002WO-US009107.
 XX PR 21-MAR-2001; 2001US-00815242.
 XX PR 06-SEP-2001; 2001US-00948993.
 XX PR 25-OCT-2001; 2001US-0342923P.
 XX PR 08-FEB-2002; 2002US-00072851.
 XX PR 06-MAR-2002; 2002US-0362659P.
 XX FA (ELIT-) ELITRA PHARM INC.
 XX FI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 FI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI: 2003-029926/02.
 XX DR N-PSDB; ACA33755.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 57809; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *X. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 669 AA;
 Query Match 66.0%; Score 2295; DB 6; Length 669;
 Best Local Similarity 65.2%; Pred. NO. 1.9e-184;
 Matches 438; Conservative 93; Mismatches 129; Indels 12; Gaps 9;

DB 488 IFNQLGSEWNLEGLHGFPPSGTKVAVGTPFPRLDMEEVAFTQEKMSGQTQNE 547
 QY 538 FAVEKEMNPDEVELKLNKD-EIKFDFDKVEIRVAEVKEVSKVSGDKLQFRLDAGDE 596
 DB 548 DTV--KMDPEETELVSTKEQIKFDPFEKVKELKVAEVINCQKVEGADKLLQFRLDAGDSQ 605
 QY 597 DROILSGIAKYPNQBELVGVKVOIVANLKRKMKKVVSGMILSAB-HDGKLTLLTVD 655
 DB 606 DROILSGIAEFYDPDSLEIGKVKWIVANLKRK-MRQGISQGMILSABAPDGSLSQVIEAP 664
 QY 656 PAVPNGSVI 664
 DB 665 KSMPNGSEI 673

RESULT 14

AAU35070
 ID AAU35070 standard; protein; 677 AA.

XX AAU35070;

XX 13-FEB-2002 (first entry)

XX Enterococcus faecalis cellular proliferation protein #357.

XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.

XX Enterococcus faecalis.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

XX 23-MAY-2000; 2000US-0206848P.

XX 26-MAY-2000; 2000US-0207727P.

XX 23-OCT-2000; 2000US-0242578P.

XX 27-NOV-2000; 2000US-0253625P.

XX 22-DEC-2000; 2000US-0257531P.

XX 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX N-PSDB; AAS52929.

XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.

XX Example 3; SEQ ID NO 10663; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes,
 CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
 CC coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
 CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 677 AA;

Query Match 65.8%; Score 2288; DB 4; Length 677;
 Best Local Similarity 63.9%; Pred. No. 7.7e-184;
 Matches 428; Conservative 104; Mismatches 130; Indels 8; Gaps 7;

QY 1 MSEK-NFYITTTIYPPSGKLHIGSATYTTIACDVLARYKELMGYDVFYLTGLDHEGQKIQ 59
 DB 9 MSQKETFYITTTIYPPSGKLHIGSATYTTIACDVLARYKELMGYDVFYLTGLDHEGQKIEK 68
 QY 60 KAEERGITPOAYVDGMVGVKELWQLDISYDKFTRTTDDYHEKVVAVQVFERLLAQDDIY 119
 DB 69 KAAELNVTPEVVDKMAADVQKLWLTLDISYDKFTRTTDDYHMAAVQQIFDELVEGSDIY 128
 QY 120 LGSEYSGWYSVDEEFTTESQLAEVFRDAGNVTGCIAPSGHEVWVSSESYFLRLSKYQD 179
 DB 129 LGSEYSGWYSVDEEFTTESQLAEVFRDAGNVTGCIAPSGHEVWVSSESYFLRLSKYQD 188
 QY 180 RLVEPFKAHPEITPDGRNLNEMLRNFIERGELEDLAVSRRTFTMGVPSNPKNHVVYVWD 239
 DB 189 RLQVYEEHPEFIQPSRKNEMLNFIKPLGLEDLAVSRRTFTMGVPSNPKNHVVYVWD 248
 QY 240 ALINAYATALGYAQDEHGNFDFWNGTVFHMVGKDLIRFHSIYWPILLMMLDVKLPDLRIA 299
 DB 249 ALSNYITALGYGSEDDSLFKYWPANV-QMWKEIVRPHTIYWPIMLMDLPLPKVFG 307
 QY 300 HGWFWKMGKSKSGNVVYPELVRYGLDPLRYVLMENLPGVSDGTFTTPTDYGRINY 359
 DB 308 HGWLLMKDGKSKSGNVVYPELVRYGLDALRYLLRAIFGSDGVFTPEDFVSRNLNY 367
 QY 360 ELANDLGNLILNRTVSMINKYFDGQIPAYVEGVTEDFHVLAEVAEKSIAADFHTHMEAVDYP 419
 DB 368 DLANDLGNLILNRTIAMINKYCDGKVPAYASKYTPFDSLSLTTAANVIGKYHEAMEKVEFN 427
 QY 420 RALEAVWTLISRTNKYIDTAPWVLDKDEALDQASVMSHWQASIRVVAHLIEFPMET 479
 DB 428 TATAEITWLVSRANKYIDTAPWVLAKEEKEKESESVMIHLAELRIVAILLOPMTET 487
 QY 480 SRAVLTOIGLE-EVSSLENLSLADPADVTVVAKGTPFPRLNMEELIAYIKQEGKNKP 538
 DB 488 PGKIFQLGLDPTMNMENIHFGEPDVTVTSKGTPIFPRLEITEVTYTIQKMSQSES 547
 QY 539 AVEKE--WNPDEVELKLNKD-EIKFDFDKVIRVAEVAEKSKEVSGDKLLQFRLDAGD 595
 DB 548 ATEEDIKANPEETTLVSTKEKQIKYDDDFDKVELKVAEVIDCKKVGADKLLQFRLDAGDE 607
 QY 596 EDROILSGIAKYFNEQELVGVKVOIVANLKRKMKKVVSGMILSABH-DGKLTLLTV 654
 DB 608 NHRQILSGIAEFYDPDPAALIGKVKWIVANLKRK-MRQGISQGMILSABSPGKQIVRA 666
 QY 655 DPAVNGSVI 664
 DB 667 KEMPNGAGI 676

RESULT 15

ABU14551
 ID ABU14551 standard; protein; 677 AA.

XX ABU14551;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #78.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Enterococcus faecalis.

XX WO200277183-A2.

XX 03-OCT-2002.
XX 21-MAR-2002; 2002WQ-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 26-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haeelbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ACA18421.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 42475; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 677 AA;
SQ Query Match 65.8%; Score 2288; DB 6; Length 677;
Best Local Similarity 63.9%; Pred. No. 7.7e-184;
Matches 428; Conservative 104; Mismatches 130; Indels 8; Gaps 7;
QY 1 MSEK-NFYITTTIYPYSGKLHIGSAITTIACDLVARYKELMGYDYFYLTLGLDEHGQKIQ 59
DB 9 MSQKETFYITTTIYPYSGKLHIGSNYTTIACDAMARYKELMGFDVYLTGVDEHGQKIEK 68
QY 60 KAEAGITQAYVDGMVAGVKELWQLLDISYDKFRTTDDYHEKVVAQVFERLLAODDIY 119
DB 69 KAAELNVTTKEYYDKWAARDVQKLWKTLDISYDKFRTTDDYEMAAVQQIFDLVEQGDY 128
QY 120 LGEYSQWYSDBEFTTESQAEVFRDEAGNTGGIAPSGHEVWVSEYFRLRSKYQD 179

DB 129 LGEYEGWYSVDSBEFTTETQAEVYRDDGNTGGKAPSGHEVELVKEESYFFRMSKYAD 188
QY 180 RLVEFFKAHPFITPDGRLEMLNRTFEGLEDLAVSRRTFTTWGVPVPSPKVVVYVWID 239
DB 189 RLVOYEEHPEFIQPSERKNEMLNNFIKPGLEDLAVSRRTFTSWGIPUKNDPKHVVVWID 248
QY 240 ALLNYATALGYAQDEHGNFDFWNGTVFHVAGKDIILRFHSIYWPILLMLMDLKLDELIA 299
DB 249 ALSNYITATLGYGSEDDSLFQKYPANV-QMVGEIVRFHTIYWPIMMALDLDLPKKVFG 307
QY 300 HGWFVMDKGMXSKSGNVVYPEMLVERYSGLDPLRYLYMLRNLPGVSGDTFTFPEDYVGRIN 359
DB 308 HGWLLXKDGXSKSGNVVYPEMLVERYSGLDPLRYLYMLRNLPGVSGDTFTFPEDYVGRIN 367
QY 360 ELANDLGNLNRVTVMINKYFDGQIPAYVEGVTEFDFHVLAEVAKSIADPHTHNEAVDYP 419
DB 368 DLANDLGNLNRVTAMINKYCDQKVPAYASKVTPFDSSELSTTAANVTGKYHEAMEKMEFN 427
QY 420 RALEAVTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLTEPEMNET 479
DB 428 TAIAEITLVSANKYIDETAPWVLAKEEKRELESVMHLAESLRIVAILLOPVMTET 487
QY 480 SRAVLTQIGLE-EVSSLLENLSLADFFADVTVAAGTPIFFRLNMEEBEIAIYKQMEGNKP 538
DB 488 PGKIFEQLGLDPEMTNMNIHFGEFTDVTVTSGTPIFFRLTEIETEVTYIKKMSQSES 547
QY 539 AVEKE--WNPDEVELKLNKD-EIKFEDFDKVEIRVAEVEKSVKVEGSDKLLQRLDAGDG 595
DB 548 ATESDIKWNPEETTLVSTKQKQIKYDDFDKVELKVAEVIDCKKVKGADKLLQRLDAGDE 607
QY 596 EDRQILSGIARYPNEQELVGKVKQIVANLKPMMKKYVSQGMILSAEH-DGKLTLLTV 654
DB 608 NHRQILSGIAEFYDPAALIGKKVIVANLKPRLK-HEGQISQGMILSAESPEGKLQIVEA 666
QY 655 DPVDPNGSVI 664
DB 667 PKENPENGAGI 676

Search completed: December 17, 2004, 21:56:06
Job time : 164 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2004, 21:51:16 ; Search time 40 Seconds
(without alignments)
1102.537 Million cell updates/sec

Title: US-10-606-607-2

Perfect score: 3479

Sequence: 1 MSEKNFYITPIYPSGKLH.....DGKLTLLTVDPAPNGSVIG 665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/aaa/5A-COMB.pep:*
- 2: /cgn2_6/prodata/1/aaa/5B-COMB.pep:*
- 3: /cgn2_6/prodata/1/aaa/6A-COMB.pep:*
- 4: /cgn2_6/prodata/1/aaa/6B-COMB.pep:*
- 5: /cgn2_6/prodata/1/aaa/6C-COMB.pep:*
- 6: /cgn2_6/prodata/1/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3479	100.0	665	3	US-08-844-059-2
2	3479	100.0	665	3	US-09-431-202-2
3	3429	98.6	679	4	US-09-583-110-2751
4	2291.5	65.9	674	4	US-09-107-532A-6201
5	1922.5	55.3	660	3	US-09-134-001C-3350
6	1767.5	50.8	618	4	US-09-710-279-860
7	1492.5	42.9	493	4	US-09-134-000C-4034
8	1137.5	32.7	648	1	US-08-451-715A-4
9	1033	29.7	521	1	US-08-584-226-2
10	755	21.7	683	4	US-09-543-681A-5403
11	726	20.9	716	4	US-09-489-039A-8243
12	701.5	20.2	402	4	US-09-248-796A-16768
13	661	19.0	720	4	US-09-328-352-4765
14	633.5	18.2	690	4	US-09-540-236-3507
15	625	18.0	680	4	US-09-252-991A-17566
16	619	17.8	382	4	US-09-270-767-62086
17	614	17.6	127	3	US-08-858-207A-490
18	574	16.5	115	3	US-08-844-059-4
19	574	16.5	115	3	US-09-431-202-4
20	493	14.2	554	4	US-09-198-452A-140
21	470.5	13.5	659	3	US-09-392-772-10
22	431.5	12.4	176	4	US-09-710-279-2332
23	344	9.0	152	3	US-09-392-772-12
24	310	8.9	888	3	US-09-134-001C-3032
25	308.5	8.9	876	1	US-08-785-429-2
26	308.5	8.9	876	3	US-08-996-621-2
27	300	8.6	883	4	US-09-583-110-2900

28 299 8.6 883 2 US-08-953-492-2 Sequence 2, Appli
29 296 8.5 902 4 US-09-107-532A-3837 Sequence 3837, Ap
30 286 8.2 894 4 US-09-134-000C-5731 Sequence 5731, Ap
31 265.5 7.6 80 4 US-09-134-000C-3795 Sequence 3795, Ap
32 256.5 7.4 804 1 US-08-785-428-2 Sequence 2, Appli
33 256.5 7.4 804 2 US-08-996-797-2 Sequence 2, Appli
34 240 6.9 128 4 US-09-270-767-62086 Sequence 62086, A
35 236 6.8 872 1 US-08-451-715A-8 Sequence 8, Appli
36 233.5 6.7 835 3 US-09-134-001C-5105 Sequence 5105, Ap
37 236.5 6.5 1045 1 US-08-452-083-2 Sequence 2, Appli
38 220.5 6.3 899 3 US-09-413-814-5 Sequence 5, Appli
39 216 6.2 318 4 US-09-248-796A-18266 Sequence 18266, A
40 213 6.1 824 4 US-09-134-000C-4908 Sequence 4908, Ap
41 212.5 6.1 833 2 US-08-844-086-2 Sequence 2, Appli
42 212.5 6.1 833 3 US-09-018-211-2 Sequence 2, Appli
43 212 6.1 804 3 US-08-855-910-4 Sequence 4, Appli
44 210.5 6.1 833 4 US-09-583-110-4098 Sequence 4098, Ap
45 205.5 5.9 828 4 US-09-107-532A-4191 Sequence 4191, Ap

ALIGNMENTS

RESULT 1
US-08-844-059-2
; Sequence 2, Application US/08844059
; Patent No. 6001601
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: NO. 6001601el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,059
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9607999.1
; FILING DATE: 16-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 665 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-844-059-2

Query Match 100.0%; Score 3479; DB 3; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.6e-303;
Matches 665; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
1 MSEKNFYITPIYPSGKLHIGSAYTTIACDVLARYKRLMGVDVYLTGLDHGQKIQK 60
|||||

OY

```

Db 1 MSEKRFYITPIYPSGKLGHSAYTTIACDVLARYKRLMGVDYFYLTLGLDEHGQKIQOK 60
QY 61 ABEAGITPOAYVDGMAGVKELWQLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYL 120
Db 61 ABEAGITPOAYVDGMAGVKELWQLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYL 120
QY 121 GEYSWYSVDEEFTESQLAVERDEAGNTGGIAPSGHEVWVSSEYFLRLSKYQDR 180
Db 121 GEYSWYSVDEEFTESQLAVERDEAGNTGGIAPSGHEVWVSSEYFLRLSKYQDR 180
QY 181 LVEFFKAHPEFITPDGRNLEMLRNFIEPGLDLAVSRITFTWGPVPSNPKHVVYVWIDA 240
Db 181 LVEFFKAHPEFITPDGRNLEMLRNFIEPGLDLAVSRITFTWGPVPSNPKHVVYVWIDA 240
QY 241 LLYATATGYAQDEHGNFKWNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLRIAH 300
Db 241 LLYATATGYAQDEHGNFKWNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLRIAH 300
QY 301 GWFVMDGKMSKSGNVVYPEMLVERYGLDPLRYLLMRNLFPVSGDGTFTPDYVGRINYE 360
Db 301 GWFVMDGKMSKSGNVVYPEMLVERYGLDPLRYLLMRNLFPVSGDGTFTPDYVGRINYE 360
QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADFTHTHEAVDPR 420
Db 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADFTHTHEAVDPR 420
QY 421 ALEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVMSHWQASIRVVVAHLIEPFMMETS 480
Db 421 ALEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVMSHWQASIRVVVAHLIEPFMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPRLNMEEBIAYIKQMEGNKPAV 540
Db 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPRLNMEEBIAYIKQMEGNKPAV 540
QY 541 EKWNPDVELKLNKDEIKFDPKVEIRVAEVEKSVKESGDKLQRLDAGGEDRQI 600
Db 541 EKWNPDVELKLNKDEIKFDPKVEIRVAEVEKSVKESGDKLQRLDAGGEDRQI 600
QY 601 LSGIAKYYPNEQELVGKVKQIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVPN 660
Db 601 LSGIAKYYPNEQELVGKVKQIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVPN 660
QY 661 GSVIG 665
Db 661 GSVIG 665

```

RESULT 2

```

US-09-431-202-2
; Sequence 2, Application US/09431202
; Patent No. 6294175
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6294175el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSES: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,202
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,059

```

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Giamli, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 665 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-431-202-2

Query Match 100.0%; Score 3479; DB 3; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.6e-303;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSEKRFYITPIYPSGKLGHSAYTTIACDVLARYKRLMGVDYFYLTLGLDEHGQKIQOK 60
Db 1 MSEKRFYITPIYPSGKLGHSAYTTIACDVLARYKRLMGVDYFYLTLGLDEHGQKIQOK 60
QY 61 ABEAGITPOAYVDGMAGVKELWQLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYL 120
Db 61 ABEAGITPOAYVDGMAGVKELWQLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYL 120
QY 121 GEYSWYSVDEEFTESQLAVERDEAGNTGGIAPSGHEVWVSSEYFLRLSKYQDR 180
Db 121 GEYSWYSVDEEFTESQLAVERDEAGNTGGIAPSGHEVWVSSEYFLRLSKYQDR 180
QY 181 LVEFFKAHPEFITPDGRNLEMLRNFIEPGLDLAVSRITFTWGPVPSNPKHVVYVWIDA 240
Db 181 LVEFFKAHPEFITPDGRNLEMLRNFIEPGLDLAVSRITFTWGPVPSNPKHVVYVWIDA 240
QY 241 LLYATATGYAQDEHGNFKWNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLRIAH 300
Db 241 LLYATATGYAQDEHGNFKWNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLRIAH 300
QY 301 GWFVMDGKMSKSGNVVYPEMLVERYGLDPLRYLLMRNLFPVSGDGTFTPDYVGRINYE 360
Db 301 GWFVMDGKMSKSGNVVYPEMLVERYGLDPLRYLLMRNLFPVSGDGTFTPDYVGRINYE 360
QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADFTHTHEAVDPR 420
Db 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADFTHTHEAVDPR 420
QY 421 ALEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVMSHWQASIRVVVAHLIEPFMMETS 480
Db 421 ALEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVMSHWQASIRVVVAHLIEPFMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPRLNMEEBIAYIKQMEGNKPAV 540
Db 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPRLNMEEBIAYIKQMEGNKPAV 540
QY 541 EKWNPDVELKLNKDEIKFDPKVEIRVAEVEKSVKESGDKLQRLDAGGEDRQI 600
Db 541 EKWNPDVELKLNKDEIKFDPKVEIRVAEVEKSVKESGDKLQRLDAGGEDRQI 600
QY 601 LSGIAKYYPNEQELVGKVKQIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVPN 660
Db 601 LSGIAKYYPNEQELVGKVKQIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVPN 660
QY 661 GSVIG 665
Db 661 GSVIG 665

```

RESULT 3

US-09-583-110-2751

; Sequence 2751, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 05/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2751
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2751

Query Match 98.6%; Score 3429; DB 4; Length 679;
Best Local Similarity 98.8%; Pred. No. 5,1e-299;
Matches 657; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy	1	MSEKNFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVYLTGLDEHGQKIQQK	60
Db	15	MSEKNFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVYLTGLDEHGQKIQQK	74
Qy	61	ABEAGITPOAVDGMVGVKELWQLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL	120
Db	75	ABEAGITPOAVDGMVGVKELWQLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL	134
Qy	121	GEYSGWTSVSDDEEFTTESQLAEVFRDEAGNVTGGIAPSGHEVWVSSESYFLRLSKYQDR	180
Db	135	GEYSGWTSVSDDEEFTTESQLAEVFRDEAGNVTGGIAPSGHEVWVSSESYFLRLSKYQDR	194
Qy	181	LVEFFKAHPFETPDGRLNEMLRNFIEPGLDLAVSRITTTTGWVPSPNKHVYVWIDA	240
Db	195	LVEFFKAHPFETPDGRLNEMLRNFIEPGLDLAVSRITTTTGWVPSPNKHVYVWIDA	254
Qy	241	LLNVTATGVAQDEHGFNDFWNGTVFHMVKDILRPHSIYWPILLMLLDVLPDLRIAH	300
Db	255	LLNVTATGVAQDEHGFNDFWNGTVFHMVKDILRPHSIYWPILLMLLDVLPDLRIAH	314
Qy	301	GWFMKDGKSKSGNYYVPEMLVERYGCLDPLRYLLMRNLPVSGDGTFTPEDYVGRINYE	360
Db	315	GWFMKDGKSKSGNYYVPEMLVERYGCLDPLRYLLMRNLPVSGDGTFTPEDYVGRINYE	374
Qy	361	LANDGLNLRNTSMINKYFDGQIPAYVEGVTBFDHVAEVAEKSIADEHTEAVDYPR	420
Db	375	LANDFGNLRNTSMINKYFDGQIPAYVEGVTBFDHVAEVAEKSIADEHTEAVDYPR	434
Qy	421	ALEAVMTLISRTNKYIDETAPVLDKDEALRDQLASVMSHWQASIRVAHLIEPFWMEIS	480
Db	435	ALEAVMTLISRTNKYIDETAPVLDKDEALRDQLASVMSHWQASIRVAHLIEPFWMEIS	494
Qy	481	RAVLTQGLBEVSLNLSADPADVTVAKGTPIFPLNMBEETAYIKEQVEGNKPAV	540
Db	495	RAVLTQGLBEVSLNLSADPADVTVAKGTPIFPLNMBEETAYIKEQVEGNKPAV	554
Qy	541	EKEWNPDEVELKLNKDKIEKPEDFKVEIRVAEVEKSVKSGSKLQFLDAGDGEDROI	600
Db	555	EKEWNPDEVELKLNKDKIEKPEDFKVEIRVAEVEKSVKSGSKLQFLDAGDGEDROI	614
Qy	601	LSGIATKYPNEQELVGKKVQIVANLKPQKMKKYVSGMILSAEHGKTLTLLTVDPAVN	660
Db	615	LSGIATKYPNEQELVGKKVQIVANLKPQKMKKYVSGMILSAEHGKTLTLLTVDPAVN	674
Qy	661	GSVIG 665	
Db	675	GSVIG 679	

RESULT 4

US-09-107-532A-6201
; Sequence 6201, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arintello, Pamela Denke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6201:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 674 amino acids
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...674
; SEQUENCE DESCRIPTION: SEQ ID NO: 6201:
US-09-107-532A-6201

Query Match 65.9%; Score 2291.5; DB 4; Length 674;
Best Local Similarity 65.0%; Pred. No. 7,5e-197;
Matches 435; Conservative 94; Mismatches 129; Indels 11; Gaps 8;

Qy	3	EKNFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVYLTGLDEHGQKIQQKAE	62
Db	9	KETFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVYLTGLDEHGQKIQQKAE	68
Qy	63	EAGITPOAVDGMVGVKELWQLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYLGE	122
Db	69	ELGVKPEYVDKMAADVKKMLKTLDISYDKFIRTTDDYHKAQVKIFRLEQGDIIYLGE	128
Qy	123	YSGWTSVSDDEEFTTESQLAEVFRDEAGNVTGGIAPSGHEVWVSSESYFLRLSKYQDRIV	182
Db	129	YEGWTSVSDDEEFTTESQLAEVYKDEGNVIGKAPSGCHEVELKEESYFFRMSKYADRL	188
Qy	183	BFEKAHPFETPDGRLNEMLRNFIEPGLDLAVSRITTTTGWVPSPNKHVYVWIDALL	242
Db	189	EYEEHPFETPDGRLNEMLRNFIEPGLDLAVSRITTTTGWVPSPNKHVYVWIDALLS	248

243 NYATGALGAQDEHGNDFKFWNGTVPFHMVGKDIILRFHSIYWPILLMMLDVKLPRDLIAHG 302
249 NYITAGYSEDESFEKIWPADV-QMVGKEIVRHTIYWPIMLMALDLPFKKVFHG 307
303 FVMKDGKMSKSGNVVYPEMLVERGLDPLRYLLMRNLPGVSDGTFPTEDVYGRIN 362
308 LLMKDGKMSKSGNVVYPEMLVERGLDPLRYLLMRNLPGVSDGTFPTEDVYGRIN 367
363 NDLGNLNTVSMINKYFDGQIPAYVEGTEPDHVLAEVAEKSIADFTHTHEAVDYPAL 422
368 NDLGNLNTVSMINKYFDGQIPAYVEGTEPDHVLAEVAEKSIADFTHTHEAVDYPAL 427
423 EAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLBPFFMETSRA 482
428 AETWLVISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLBPFFMETSRA 487
483 VLTQIGLE-EVSSLENLSLADPPADVTVVAKGTPFPRLNMEETIAYIKQM-EG---NK 537
488 IFNLGLDSETNLEGLHGFEPGSGTKVAKGTPFPRLNMEETIAYIKQM-EG---NK 547
538 PAVEKEWNPDEVELKLNKD-EIKFEDFDKVEIRVAEVEKSVKVEGSDKLLQFRLDAGDGE 596
548 DTV--KWDPEEELVSTKQIKFDFPERKVELKVAEINQCKVEGADKLLQFRLDAGDGE 605
597 DRQILSGIAKYPNEQELVCKKQVIVANLKPRMMKKYVSQGMILSAB-HDCKLTLTVD 655
606 DRQILSGIAEFYDPSELIGKVKVIVANLKPRK-MRGQISOGMILSABEPDGSLOVIEAP 664
656 PAVPNGSVI 664
665 KSPNGSEI 673

RESULT 5

US-09-134-001C-3350
; Sequence 3350, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3350
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3350

Query Match 55.3%; Score 1922.5; DB 3; Length 660;
Best Local Similarity 55.9%; Pred. No. 1e-163;
Matches 379; Conservative 113; Mismatches 160; Indels 23; Gaps 11;
QY 1 MSEKNFYITPIYPSGKLHIGSAITACDLARYKRLMGYDVFYLTGLDHEGKIQOK 60
DB 5 MAKDTFYITPIYPSGKLHIGSAITACDLARYKRLMGYDVFYLTGLDHEGKIQOK 64
QY 61 ABEAGITPQAYVDMGAMGVKELQLLDISYDKFRTTDDYHEKVAQVFERLLAODDIYL 120
DB 65 AQAAGKTELEYLDLDMISGINKLMSKLSISNDFFIRITEERHKQVKEVFERLLKQSDIYL 124
QY 121 GEYSGMTSVSDEEFTTESQLAE-VFRDEAGNVTGGIAP-SGHEVWVSEESYFLRLSKYQ 178
DB 125 GEYSGMTSVDPETTYTESQLVDPVY--ENKIVGKSPDSCHEVELVKEESYFFNINKYT 182
QY 179 DLVEFFKAPHEFTTPDGRNEMLRNTEIEQLEDLAVSRFTTTPGVPSPNPKHVVYVWI 238

183 DRLEFYDENPFIQPPSRKNEMINNFIRKPGLEDLAVSRFTTTPGVPSPNPKHVVYVWI 242
239 DALLNYATGALGAQDEHGNDFKFWNGTVPFHMVGKDIILRFHSIYWPILLMMLDVKLPRDLI 298
243 DALVNYISSIYGLSDDETLFNKYWPADI-HLMAKEIVRPHSIWPIILLMALDLPFKKVF 301
299 AHGFWKMGKMSKSGNVVYPEMLVERGLDPLRYLLMRNLPGVSDGTFPTEDVYGRIN 358
302 AHGFWKMGKMSKSGNVVYPEMLVERGLDPLRYLLMRNLPGVSDGTFPTEDVYGRIN 361
359 YELANDLGNLNTVSMINKYFDGQIPAYVEGTEPDHVLAEVAEKSIADFTHTHEAVD 418
362 YELANDLGNLNTVSMINKYFDGQIPAYVEGTEPDHVLAEVAEKSIADFTHTHEAVD 421
419 PRALAEVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLBPFFMETSRA 478
422 SVALSTWKFISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLBPFFMETSRA 481
479 TSRAVLTQIG-----LEEVSSLENLSLADPPADVTVVAKGTPFPRLNMEETIAYIKQM 533
482 APREIFKQLNNPDDLHQLDSLOQYGM--LSEAITVTEKTPFPRLDTEAEIAYIKESM 539
534 EGNKPAVEKEWNPDEVELKLNKD-EIKFEDFDKVEIRVAEVEKSVKVEGSDKLLQFRLDAG 593
540 QPKK-SIKQSDPE-----GKEQIDIKDFKVEIKAAITIDAENVKSEKLLKIKVEL- 590
594 DGEDRQILSGIAKYPNEQELVCKKQVIVANLKPRMMKKYVSQGMILSAB-HDCKLTLT 653
591 DNEQRQIVSGIAKYPNEQELVCKKQVIVANLKPRMMKKYVSQGMILSAB-HDCKLTLT 648
654 VDPAPNGSVI 664
649 LPSAIPNGSVI 659

RESULT 6

US-09-710-279-860
; Sequence 860, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 860
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-860

Query Match 50.8%; Score 1767.5; DB 4; Length 618;
Best Local Similarity 54.8%; Pred. No. 7.7e-150;
Matches 347; Conservative 108; Mismatches 155; Indels 23; Gaps 11;
QY 39 LMGYDVFYLTGLDHEGKIQOKAEAGITPQAYVDMGAMGVKELWQLLDISYDKFRTTDD 98
DB 1 MOGYDVFYLTGLDHEGKIQOKAEAGITPQAYVDMGAMGVKELWQLLDISYDKFRTTDD 60
QY 99 DYHEKVAQVFERLLAODDIYLGYSYGSYSVDEEFTTESQLAE-VFRDEAGNVTGGIAP 157
DB 61 ERHKQVKEVFERLLKQSDIYLGYSYGSYSVDEEFTTESQLAE-VFRDEAGNVTGGIAP 118
QY 158 -SGHEVWVSEESYFLRLSKYQDRLVEFFKAPHEFTTPDGRNEMLRNTEIEQLEDLAVSRFTTTPGVPSPNPKHVVYVWI 216
DB 119 DSGHEVELVKEESYFFNINKYTDRLEFFYDENPFIQPPSRKNEMINNFIRKPGLEDLAVSRFTTTPGVPSPNPKHVVYVWI 178

QY 217 RTTWTGVPSPKHHVYVWIDALLNATAGYAQDEHGNEDKFWNGTGFHVMGKDLR 276
DB 179 RTSFDWGRVPSNPKHHVYVWIDALVNISSLGYSDETLFNKYWPADI-HLMAKEIVR 237
QY 277 FHSIYWPILLMLDVKLDPDLIAHGWFMKGKSKGNVYVPEMLVRYGLDPLRYYL 336
DB 238 FHSIYWPILLMALDPLPKVFAHGWFMKGKSKGNVYVPEMLVRYGLDPLRYYL 297
QY 337 MRNLPVSGDGTFTPEDYVGRINVELANDLGNLNRVTSMINKYVFDGOIPAYVSGVTEFDH 396
DB 298 MRELPFGSDGVTFPEAFVETRYNDLANDLGNLNRVTSMINKYVFDGOIPAYVSGVTEFDH 357
QY 397 VLAEVAEKSIADFRTHMEAVDYPRALEAVMTLISRTNKYIDETAPWYLDKDEALRDQLAS 456
DB 358 KMEAMALTVKSFNDNMESLQFSVALSTVMKFSRTNKYIDETQWPWLAKDENQREMLGN 417
QY 457 VMSHWQASIRVVAHLIEPFWMETSRVLTQIG-----LEEVSSLENLSLADFPADVTVVA 511
DB 418 VMAHLVNERIFATLLQFFTHAPREIFKQINNNPDLHQLDLSLQOYGM--LSEAITVTE 475
QY 512 KGTFIFRLNMEEEIAYIKQMEGNKPAVEKWNPDVEVLKLNKDEIKFDFDKVEIRVA 571
DB 476 KPTIFRPLDTEAIIAYIKESMQPPK-SIKQSDP-----GKEQIDIKDFDKVEIKAA 527
QY 572 EVKEVSKVSGDKLLQRLDAGDGEDRQLLSGIAYKYPNEQELVKKVQIVANLKPRKOM 631
DB 528 TIIDAENVKSEKLLIKVEL-DNEQRQIVSGIAKFY-RPEDIIGKRVAVVTLNPKAKLM 585
QY 632 KKYVSOQMLSAHDGKLTLLTVDPAVPGNSVI 664
DB 586 GQ-KSEGMLSAEKGVTLLISLPSAIPNGAVI 617

RESULT 7

US-09-134-000C-4034
; Sequence 4034, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4034
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4034

Query Match 42.9%; Score 1492.5; DB 4; Length 493;
Best Local Similarity 60.1%; Pred. No. 2.7e-125;
Matches 283; Conservative 79; Mismatches 102; Indels 7; Gaps 6;
QY 199 NEMLRNTEPGLDLAVSRRTFTTGVPSNPHVTVWIDALNATAGYAQDEHGNF 258
DB 24 NEMLRNTEPGLDLAVSRRTFTTGVPSNPHVTVWIDALNATAGYAQDEHGNF 83
QY 259 DKFWNGTGFHVMGKDLRHSIYWPILLMLDVKLDPRLIAHGWFMKGKSKGNVY 318
DB 84 QKYPANV-QVNGKEIVRFHTIYWPIMMALDPLPKVFGHLLMKOGRSKGNVY 142
QY 319 YPEMLVRYGLDPLRYILMRNLVPGSDGTFTPEDYVGRINVELANDLGNLNRVTSMINK 378
DB 143 YPEMLVRYGLDALRYLLRAIPFGSDGVTFPEDYVGRINVELANDLGNLNRVTSMINK 202
QY 379 YFDQIQIPAYVSGVTEFDHVAEAKSIADFRTHMEAVDYPRALEAVMTLISRTNKYIDE 438

DB 203 YCDGKVPAYASKVTFFDSELSTTAANVIGKVHEAMEKMEFNATIAEITWLVSRANKYIDE 262
QY 439 TAPWYLDKDEALRDQLASVMSHWQASIRVVAHLIEPFWMETSRVLTQIGLE-EVSSLEN 497
DB 263 TAPWYLDKDEALRDQLASVMSHWQASIRVVAHLIEPFWMETSRVLTQIGLE-EVSSLEN 322
QY 498 LSLADFPADVTVVAKGTFFPRLNMEEEIAYIKQMEGNKPAVEKE--WNPDEVELKLNK 555
DB 323 IHFGFPTDVTVTSGTFFPRLNMEEEIAYIKQMEGNKPAVEKE--WNPDEVELKLNK 382
QY 556 D-EIKFEDPDKVEIRVAEVAEKVSKVSGDKLLQRLDAGDGEDRQLLSGIAYKYPNEQEL 614
DB 383 EQKIYDDPDKVELKVAEVIDCKVKYKADKLLQRLDAGDGEDRQLLSGIAYKYPNEQEL 442
QY 615 VGGKQIVANLKPRKOMKYYVSOQMLSAH-DGKLTLLTVDPAVPGNSVI 664
DB 443 IGKVVIVANLKPRK-MRQIISQMLSAEPEGKLOIVEAPKPNPGAGI 492

RESULT 8

US-08-451-715A-4
; Sequence 4, Application US/08451715A
; Patent No. 5801013
; GENERAL INFORMATION:
; APPLICANT: Tao, Jianshi
; APPLICANT: Qul, Yan
; APPLICANT: Houman, Fariba
; APPLICANT: Shen, Xiaoyu
; APPLICANT: Schimmel, Paul R.
; TITLE OF INVENTION: Helicobacter Aminoacyl-tRNA Synthetase
; TITLE OF INVENTION: Proteins, Nucleic Acids and Strains Comprising Same
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,715A
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: CPI94-25
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 648 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-451-715A-4

Query Match 32.7%; Score 1137.5; DB 1; Length 648;
Best Local Similarity 38.9%; Pred. No. 3.2e-93;
Matches 266; Conservative 117; Mismatches 233; Indels 67; Gaps 16;
QY 8 ITTPIYPSGKLHIGSAVYTIACDVLARYKRLMGVDVFLTGLDGHGKIQKAEAGIT 67
DB 6 ITTPIYVNDIPHIGHAVTTLIADTLKKYVTLQGEVFFLTGTDHGGKIEQASLRNQS 65
QY 68 POAYVGNVAGVKEIWLQDLIDISYDKFIETDDYHEKVAQVPERLLAODDIYLGYSWY 127

Db 66 PRAYADISAIKFDQWDFNLDYDGFIRTTDGEHQKVCQNAPEIMFKEGDIYKGAISGY 125
 QY 128 SVSDEFFTESQLAEVFRDEAGNVGGIAPSG-HEVEWVSEESYELRLSKYQDRLEVEPK 186
 Db 126 CVSCSYCAISK-ADNTSKV-----LCPDCLRETALLEESYFFKSAYEKPLEFYA 178
 QY 187 APEFTTPDGRNLEMRNIEPGLDLAVSRITFTWGVVPV---SNPKHVVVVWIDALLN 243
 Db 179 KNEPALPIYRKNB-VTSIEQGLDLSITRISFENGIPLPKKNMDPKHVVVVWIDALLN 237
 QY 244 YATAGYADQEHGNDKFWNGTVFHVHVKDILRFHSIYWPILLMLMDVKLPDLRIAGWF 303
 Db 238 YASALGYLGNLDNKAHF--ECARHTVGVKDIIRFHAIFYWPAFLMSLNPFLFKQLCVHGM 295
 QY 304 VMKDGWMSKSGNVVPEMLVERVYGDPLRYILMRNLPVSGDTFTPDYVGRINVELAN 363
 Db 296 TIEGVZMSKSLGNVLDQAQLAMEYGEIELRYFLREVPFGQDGSKKALIERINANLNN 355
 QY 364 DLGNLLNRTVSMINKYFD-----GQIPAYVEGVTEPDHVLAEVAKSIADFTTHMEAVDY 418
 Db 356 DLGNLLNRLGLMAKYNFYSLKSAKITAYYSKELEKAHOILDNANSFVPMQLH----- 409
 QY 419 PRAEAVWTLISFTKYIDETAPVWLDKDEALRDQLASVMHQAQSIIRVVAHLIEPFWME 478
 Db 410 -KALELFNVYDFINKLIAKEEPWLVKKNB-SEKLEALLSLIANALLOSSFLLYAFMPK 467
 QY 479 TSRAVLTQIGLE-----EVSSLENLSLADPPADVTVVAKGTPFPRLNMBEEIAY 528
 Db 468 SAVKLASAFNTEITPNRYERFFKAKQLQDWLQD-----TEPLFCM----- 509
 QY 529 IKQMEGNPAPVKEWNPDEVELK-----LNKBIKEDFDKVEIRVAEVEKSKVEG 581
 Db 510 --EKIEKTDKR-BKEYPPEKAKEKEKAPPKQENYIGIEDFKVEIKVGLIKEAQRIEK 566
 QY 582 SDKLLQFRIDAGGEDRQILSGTAKYYPNEQELVKGKVOIVANLKPRKMKKYVSGMIL 641
 Db 567 SNKLLRLKVDLGNRLRQIISGLALDYEPB-SLVGQWCVVANLKPALMGB-MSEGMIL 624
 QY 642 SAHDGKLLTLLVDPAVNGSVI 664
 Db 625 AVRDSDNALISPTKEKIAGSLI 647

RESULT 9

US-08-584-226-2
 ; Sequence 2, Application US/08584226
 ; Patent No. 5798240
 ; GENERAL INFORMATION:
 ; APPLICANT: Martinis, Susan A.
 ; APPLICANT: Sasanfar, Mandana
 ; APPLICANT: Kim, Sunghoon
 ; APPLICANT: Lee, Sang Ho
 ; APPLICANT: Schimmel, Paul R.
 ; TITLE OF INVENTION: RECOMBINANT MYCOBACTERIAL METHIONYL-CRNA
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 ; STREET: Two Militia Drive
 ; CITY: Lexington
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02173-4799
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/584,226
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ;

APPLICATION NUMBER: US 08/305,766
 FILING DATE: 13-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Brook, David E.
 REGISTRATION NUMBER: 22,592
 REFERENCE/DOCKET NUMBER: CP194-05Z
 TELEPHONE: 617-861-6240
 TELEFAX: 617-861-9540
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 521 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-584-226-2

Query Match 29.7% Score 1033; DB 1; Length 521;
 Best Local Similarity 41.0%; Pred. No. 5,5e-84;
 Matches 214; Conservative 94; Mismatches 200; Indels 14; Gaps 7;

QY 1 MSEKNFYITTPYPSGKLHIGSAYTTIACDLVARYKRLMGYDVYLTGLDEHGKIQOK 60
 Db 1 MEMKPYVTTTATAYENAAHPVGHAYEYIATDAIRFKRLDGYDVRFLTGTDEHGLVAQA 60
 QY 61 AEEAGITPOAYVDGMVGVKELWQLDISYDKFIRTTDDYHEKVVAQVPERLLAQQDIYL 120
 Db 61 AAAGVPTAALARRNSDVFORMQBALNISDFRITRTTDDHHEASKELWRMSAAGDIYL 120
 QY 121 GEYSGWYSVSDSEFFFTBSQLAEVFRDEAGNVGTGIIAPSGHEVEWVSEESYFLSLKYQDR 180
 Db 121 DNTSGWYSVSDSEFFFTBSQLAEVFRDEAGNVGTGIIAPSGHEVEWVSEESYFLSLKYQDR 174
 QY 181 LVEFFKAPHEFTIPDGRNLEMRNIEPGLDLAVSRITFTWGVVPSPKHVVVWIDA 240
 Db 175 LLAHYHANPDIAPETRRNEVI-SFVSGGLDLSISRTSPDWGVQVPEHPDHYVWVVA 233
 QY 241 LLNVATAGVAQDEHGFDFWNGTVFHVHVKDILRFHSIYWPILLMLMDVKLPDLRIAH 300
 Db 234 LTNYLTGAGPDDTDSLEFRYWPADL-HMIAKDIIRFHAVYWPFLMSAGIELPRIFAH 292
 QY 301 GWFVMDGKMSKSGNVVPEMLVERYGLDPLRYILMRNLPVSGDGTFTPDYVGRINYE 360
 Db 293 GFLHNRGEKMSKSGNVIVDPVALAEALGVQVRYFLIREVFFGQGSYSDEAIVTRINTD 352
 QY 361 LANDGLNLLNRTVSMINKYFDGQIPAYVEGVTEPDHVLAEVAKSIADFTTHMEAVDYP 420
 Db 353 LANELGNLARRSLMSVAKNLDGRVNPGE-FADADAALLATADGLLERVGHFDAQAMHL 411
 QY 421 ALEAVWTLISRTNKYIDETAPVWLDKDEALRDQ--LASVMHQAQSIIRVVAHLIEPFWME 478
 Db 412 ALEAILMLGDKANKYFVSVQFPWLRKSEADQARFRTTLYTCVCEVVRVIAALLIQVMPE 471
 QY 479 TSRAVLTQIG-LEEVSSLENLSLADPPADVTVVAKGTPPIER 519
 Db 472 SAGKILDLGQAPNQRSFAAVGVRLLTEG--TALPPTPTGVFPER 511

RESULT 10

US-09-543-681A-5403
 ; Sequence 5403, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543,681A
 ; PRIOR FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128,706
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 5403

; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16768
; LENGTH: 402
; TYPE: PR1
; ORGANISM: Candida albicans
US-09-248-796A-16768

Query Match 20.2%; Score 701.5; DB 4; Length 402;
Best Local Similarity 38.7%; Pred. No. 2.1e-54;
Matches 147; Conservative 77; Mismatches 141; Indels 15; Gaps 6;
QY 2 SEKNFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGY-DVFLYGLDEHGQKIQQK 60
Db STKPFYITPIFYVNAAPHIGHLYSMIADTRNKWEKLNPLKESFMTGTDEHGLKIQT 78
QY 61 AEEAGITPOAYVDMGAVGVKELWQLDISYDKFIRTTDDYHEKVAQVPERLLAQQDIYL 120
Db AEKLGLEPKVLVDKVSQNSFKAEQFDVNYDRFIRTTDNDHIELVRYFNLNMEKGFYIT 138
QY 121 GEYSGWYSDESBEFFTESQLAIEVFRDEAGNVGTGGIAPSGHEVWSESYFLRLSKYQDR 180
Db 139 DTHSGWYS-SDTEFFETQIEEVVKN--GKAVKISSETKNEVYQKEITNYFFKLSMFQEQ 196
QY 181 LVFEFFKAHBEFIPDGRLEMLNFTPEGLDLAVSR--TTTGWVPVPSNPKHVVYWI 238
Db 197 LIQFLKQNPFEIKPKRYQFIKLEEDTKLPDISISRPSRLKWSIEVENDSTQKIYWF 256
QY 239 DALLNATATLGAQDEHGNFDRF-----WNGTVFHWGDKILRFHSIYWPILLMLDV 291
Db 257 DALLNVLTKPKPHGEVQDSKEVTPENSINWAT--HVIGKLIIRFCIYWFILMAAGI 314
QY 292 KLPDRIAHGWFMKDGKMSKGNVYPEMLVERYGLDPLRYLMLRNLPVSGDGTFTPE 351
Db 315 ELPKQIVHSHMLCDGFKMSKLSGNVDPMEISEYGVDPFFVFLVENSIDDDCKFS-E 373
QY 352 DVYGRINVELANDLGNLAR 371
Db 374 ELQSRSDAVLGYCNLSR 393

RESULT 13
US-09-328-352-4765
; Sequence 4765, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4765
; LENGTH: 720
; TYPE: PR1
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4765

Query Match 19.0%; Score 661; DB 4; Length 720;
Best Local Similarity 28.3%; Pred. No. 2.4e-50;
Matches 207; Conservative 115; Mismatches 296; Indels 114; Gaps 26;
QY 2 SEKNFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVFLYGLDEHGQKIQQKA 61
Db 33 SVRKILVTNALPYANGPIHGLHLYQADIWVRAMRANGHVDVTVCCADDAHGTALMLRA 92
QY 62 ERAGITPOAYV-----DGMVGVKELWQLDISYDKF-----IRTTDDYHE-- 102
Db 93 EANGISPEEQIANVQKEHIRDFDGFVPHD-----YDSTHSDANKARSTDIYIKNR 144

QY 103 -----KVAQVFE---RLLAQDDIYLG-----EYSGWYSVDSBEFFFTESQLAIEVRDE 147
Db EAGNIAVRPVTQLDFPEKGMFLSDRFIKGTCPKCKSGDQYGDSCVCGTITYNATELLNPR 204
QY 148 AGNVTGGIAPSGHEVWSESYFLRLSKYQDRILVEFFKAHBEFIPDGRLEMLNLR-- 204
Db 205 --STLSGATP-----VE-KSSDHYFFKLPNFAELQKWR-----DEGLPLSIANKLD 250
QY 205 -FIEPGLDLAVSRITFTGWVPVPSNPKHVVYWMIDALLNYATAL-GYAODEHG--NFDK 260
Db 251 EWPEAGLADWDISRDADPYFGFEIPDAPNPKYFYVWDAPIGYSSFFENYIKTKRPDLNFD 310
QY 261 FW-----NGTVFHWGDKILRFHSIYWPILLMLDVLPDLIAHGWFMKDGMSKGN 316
Db 311 FWKDSQNEVYHFGKDIDVYFHALFWPAMLEGANYRTFTGLFVNGFELTVNGQCKSKSRGT 370
QY 317 VVYPEMLVERYGLDPLRYLMLRNLPVSGDGTFTPEYVGRINVELANDLGNLNRVTSM 375
Db 371 FIKAEYTLQHLNPEYLYYFASKLSKVEDSDNLDDFVQKNSDLVGKVVNIASRCAKF 430
QY 376 INKYFDGQIPAYVEGVTEFPHVLAEV-AEKSIADFTHTHEAVDYPRLAEAVWTLSRTNK 434
Db 431 INSSFNNTLSS--TCAESDLVQSFIDAGDSIAAY--EAREFSTAIRIEMALADRANQ 484
QY 435 YIDETAPWLKDDEALRDQLASVMHQAISIRVVVAHLIEPFMMETSRVLTQIGLEEVSS 494
Db 485 YIDKKWALAKGEGQBQVLDVCSGVINLFRQAVLAPVLPTLAQQVDFLKLESF-- 542
QY 495 LENLSLADFPADVTV-----VAKGTTFPRLNMEESIAYI---KEQMEGNKP-----AVEK 542
Db 543 -----DFESRQILVSHIEIAQFQPLMQRVDPKAVAAVMDASKESIGAPAPQATKAAXK 595
QY 543 ENWPD-----EVELKLNKDEIKFDFDKVIRVAEVEKSVKSGDKLLQFRLDA 592
Db 596 EKSAEKKAAPTAVGEABI-----IGIEDFLKDLRVQAVVEAGTVEGSDKLLQUTLDV 649
QY 593 GDEGDRQILSGIARYPNEQELVGKVKQIVANLKPRMMKKYVSQGMILSAEHGDKLTL 652
Db 650 GEAEPRNVFSGIRSQYAPB-DLAKGLVVMVANLAPRK-MRFGISNGWLAAGGEGIFII 707
QY 653 TVDPAVENGSVI 664
Db 708 SPDSGAKPGDKV 719

RESULT 14
US-09-540-236-3507
; Sequence 3507, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3507
; LENGTH: 690
; TYPE: PR1
; ORGANISM: Moraxella catarrhalis
US-09-540-236-3507

Query Match 18.2%; Score 633.5; DB 4; Length 690;
Best Local Similarity 28.3%; Pred. No. 6.6e-48;
Matches 204; Conservative 121; Mismatches 296; Indels 101; Gaps 26;
QY 2 SEKNFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVFLYGLDEHGQKIQQKA 61
Db 10 TSKRILVTSALPYANGPIHGLHLYQADIWVRAMRANGHNVTVCCADDAHGTALMLKA 69
QY 62 ERAGITPOAYVDMGAVGVKELWQLDISYDK-----FIRTTDDYHEKVAQVPERLLAQQD 117

This Page Blank (uspto)

Sequence 43933, A
Sequence 13013, A
Sequence 71042, A
Sequence 5896, Ap
Sequence 52642, A
Sequence 51647, A
Sequence 53102, A
Sequence 46354, A
Sequence 54439, A
Sequence 76747, A
Sequence 11554, A
Sequence 6989, Ap
Sequence 6990, Ap
Sequence 58709, A
Sequence 62336, A
Sequence 62166, A
Sequence 63675, A
Sequence 64482, A
Sequence 13121, A
Sequence 64107, A
Sequence 63402, A
Sequence 210360, A
Sequence 163558, A
Sequence 7843, Ap
Sequence 68671, A
Sequence 10210, A
Sequence 56600, A
Sequence 75322, A
Sequence 76244, A
Sequence 13925, A
Sequence 4773, Ap
Sequence 2, Appli

14 1931 55.5 657 15 US-10-282-122A-43933
15 1930 55.5 657 9 US-09-815-242-13013
16 1922.5 55.3 656 15 US-10-282-122A-71042
17 1884 54.2 652 9 US-09-815-242-5896
18 1678 48.2 674 15 US-10-282-122A-52642
19 1640.5 47.2 644 15 US-10-282-122A-51647
20 1572 45.2 645 15 US-10-282-122A-53102
21 1383 39.8 670 15 US-10-282-122A-46354
22 1189.5 34.2 628 15 US-10-282-122A-54439
23 1153.5 33.2 521 15 US-10-282-122A-76747
24 1152.5 33.1 656 9 US-09-815-242-11554
25 1152.5 33.1 656 15 US-10-335-977-6989
26 1152.5 33.1 662 15 US-10-335-977-6990
27 1148.5 33.0 650 15 US-10-282-122A-58709
28 1045 30.0 519 15 US-10-282-122A-62336
29 1042.5 30.0 519 15 US-10-282-122A-62166
30 1040 29.9 537 15 US-10-282-122A-63675
31 1037 29.8 519 15 US-10-282-122A-64482
32 1015.5 29.2 389 9 US-09-815-242-13121
33 871.5 25.1 512 15 US-10-282-122A-64107
34 838 24.9 512 15 US-10-282-122A-63402
35 795 22.1 410 15 US-10-424-599-210360
36 794 22.8 579 16 US-10-437-963-163558
37 759 21.8 577 14 US-10-032-585-7843
38 755 21.7 675 15 US-10-282-122A-68671
39 729 21.0 677 9 US-09-815-242-10210
40 729 21.0 677 15 US-10-282-122A-56600
41 729 21.0 677 15 US-10-282-122A-75322
42 729 21.0 677 15 US-10-282-122A-76244
43 729 21.0 704 9 US-09-815-242-13925
44 726.5 20.9 593 15 US-10-108-260A-4773
45 726.5 20.9 593 15 US-10-606-608-2

ALIGNMENTS

RESULT 1
US-10-606-607-2
; Sequence 2, Application US/10606607
; Publication NO. US20040014121A1
; GENERAL INFORMATION:
; APPLICANT: LAWOR, ELIZABETH JANE
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GM10253B-3C1D1
; CURRENT APPLICATION NUMBER: US/10/606, 607
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: 10/187,641
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 10/025,189
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/432,692
; PRIOR FILING DATE: 1993-11-03
; PRIOR APPLICATION NUMBER: 08/844,056
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: GB 9607999.1
; PRIOR FILING DATE: 1996-04-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-606-607-2

Query Match 100.0%; Score 3479; DB 15; Length 665;
Best Local Similarity 100.0%; Pred. No. 56-263;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSEKNFYITPIYPYSGKLHISAYTTIACDLVRLKRLMGVDVFLTGLDHEGKIQQK 60
DB 1 MSEKNFYITPIYPYSGKLHISAYTTIACDLVRLKRLMGVDVFLTGLDHEGKIQQK 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: December 17, 2004, 21:56:12 ; Search time 147 seconds
(without alignments)
1618.777 Million cell updates/sec

Title: US-10-606-607-2
Perfect score: 3479
Sequence: 1 MSEKNFYITPIYPYSGKLHISAYTTIACDLVRLKRLMGVDVFLTGLDHEGKIQQK 665

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues
Total number of hits satisfying chosen parameters: 1589859

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PTC_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3479	100.0	665	15	US-10-606-607-2
2	3446	99.1	679	9	US-09-815-242-13418, A
3	3446	99.1	679	15	US-10-282-122A-73893, A
4	2893.5	82.9	666	15	US-10-282-122A-74415
5	2883	82.9	669	15	US-10-282-122A-72423, A
6	2295	66.0	669	15	US-10-282-122A-57809
7	2288	65.8	677	9	US-09-815-242-10663
8	2288	65.8	677	15	US-10-282-122A-42475
9	2284.5	65.7	666	9	US-09-815-242-4933, Ap
10	2113	60.7	664	15	US-10-282-122A-60385
11	1949	56.0	659	15	US-10-282-122A-71524
12	1935.5	55.6	669	15	US-10-282-122A-46123
13	1934	55.6	657	9	US-09-815-242-12135

```

Qy 61 AEEAGITPQAVYDGMVGVKELMQLDISYDKFIRTTDDYHEKVAVQVFERLLAQQDIYL 120
Db 61 AEEAGITPQAVYDGMVGVKELMQLDISYDKFIRTTDDYHEKVAVQVFERLLAQQDIYL 120
Qy 121 GEYSGWYSVDEBEFFTSQLAEVPRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQDR 180
Db 121 GEYSGWYSVDEBEFFTSQLAEVPRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQDR 180
Qy 181 LVEFFKAHPFFITPDGRLNEMLNFIPEGLDLAVSRITFTWGPVPSNPXHVYVWIDA 240
Db 181 LVEFFKAHPFFITPDGRLNEMLNFIPEGLDLAVSRITFTWGPVPSNPXHVYVWIDA 240
Qy 241 LLNYATAGVGAQDEHGNFDFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVKLPDLRIAH 300
Db 241 LLNYATAGVGAQDEHGNFDFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVKLPDLRIAH 300
Qy 301 GWFVKDKGMSKSGNVVYPEMLVERVGLDPLRYILMRNLPGVSDGTFPTPEDYVGRINYE 360
Db 301 GWFVKDKGMSKSGNVVYPEMLVERVGLDPLRYILMRNLPGVSDGTFPTPEDYVGRINYE 360
Qy 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVLAEVAKSIADPHTHMEAVDYPR 420
Db 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVLAEVAKSIADPHTHMEAVDYPR 420
Qy 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHQASIRVVAHLIEPFNMETS 480
Db 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHQASIRVVAHLIEPFNMETS 480
Qy 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
Db 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
Qy 541 EKWNPDDELKLNKDEIKFEDPKVEIRVAEYKEVSKVSGDKLLQFLDAGDGEDROI 600
Db 541 EKWNPDDELKLNKDEIKFEDPKVEIRVAEYKEVSKVSGDKLLQFLDAGDGEDROI 600
Qy 601 LSGIAKYYPNEQELVGKQVIVANLAKPRQMKKYVSQGMILSAEHGDKLTLLTVDPAVFN 660
Db 601 LSGIAKYYPNEQELVGKQVIVANLAKPRQMKKYVSQGMILSAEHGDKLTLLTVDPAVFN 660
Qy 661 GSVIG 665
Db 661 GSVIG 665

```

RESULT 2

```

US-09-815-242-13418
; Sequence 13418, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

```

```

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13418
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13418

Query Match      99.1%; Score 3446; DB 9; Length 679;
Best Local Similarity 99.2%; Pred. No. 2e-260;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSEKNFYITTTPIYVPSGKLGHSAYTTIACDVLARYKRLMGYDVYLTGLDEHGQKIQQ 60
Db 15 MSEKNFYITTTPIYVPSGKLGHSAYTTIACDVLARYKRLMGYDVYLTGLDEHGQKIQQ 74
Qy 61 AEEAGITPQAVYDGMVGVKELMQLDISYDKFIRTTDDYHEKVAVQVFERLLAQQDIYL 120
Db 75 AEEAGITPQAVYDGMVGVKELMQLDISYDKFIRTTDDYHEKVAVQVFERLLAQQDIYL 134
Qy 121 GEYSGWYSVDEBEFFTSQLAEVPRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQDR 180
Db 135 GEYSGWYSVDEBEFFTSQLAEVPRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQDR 194
Qy 181 LVEFFKAHPFFITPDGRLNEMLNFIPEGLDLAVSRITFTWGPVPSNPXHVYVWIDA 240
Db 195 LVEFFKAHPFFITPDGRLNEMLNFIPEGLDLAVSRITFTWGPVPSNPXHVYVWIDA 254
Qy 241 LLNYATAGVGAQDEHGNFDFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVKLPDLRIAH 300
Db 255 LLNYATAGVGAQDEHGNFDFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVKLPDLRIAH 314
Qy 301 GWFVKDKGMSKSGNVVYPEMLVERVGLDPLRYILMRNLPGVSDGTFPTPEDYVGRINYE 360
Db 315 GWFVKDKGMSKSGNVVYPEMLVERVGLDPLRYILMRNLPGVSDGTFPTPEDYVGRINYE 374
Qy 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVLAEVAKSIADPHTHMEAVDYPR 420
Db 375 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVLAEVAKSIADPHTHMEAVDYPR 434
Qy 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHQASIRVVAHLIEPFNMETS 480
Db 435 ALBAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHQASIRVVAHLIEPFNMETS 494
Qy 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
Db 495 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 554
Qy 541 EKWNPDDELKLNKDEIKFEDPKVEIRVAEYKEVSKVSGDKLLQFLDAGDGEDROI 600
Db 555 EKWNPDDELKLNKDEIKFEDPKVEIRVAEYKEVSKVSGDKLLQFLDAGDGEDROI 614
Qy 601 LSGIAKYYPNEQELVGKQVIVANLAKPRQMKKYVSQGMILSAEHGDKLTLLTVDPAVFN 660
Db 615 LSGIAKYYPNEQELVGKQVIVANLAKPRQMKKYVSQGMILSAEHGDKLTLLTVDPAVFN 674
Qy 661 GSVIG 665
Db 675 GSVIG 679

```

RESULT 3

```

US-10-282-122A-73893
; Sequence 73893, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl

```


APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 73893
LENGTH: 679
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73893

Query Match 99.1%; Score 3446; DB 15; Length 679;
Best Local Similarity 99.2%; Pred. No. 2e-260;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSEKNFYITTPYPSGKLHGSAYTTIACDVLARYKRLMGYDVFYLTGLDEHGQKIQK 60
Db 15 MSEKNFYITTPYPSGKLHGSAYTTIACDVLARYKRLMGYDVFYLTGLDEHGQKIQK 74
Qy 61 ABEAGITPQAYVDGMAGVKELWQLDISYDKFIRTTDDYHEKVAVQVFERLLAQDDIYL 120
Db 75 ABEAGITPQAYVDGMAGVKELWQLDISYDKFIRTTDDYHEKVAVQVFERLLAQDDIYL 134
Qy 121 GEYSGWYSVDEBFTTESQLAEVFRDEAGNVTGCIAPSGHEVWVSEESYFLRLSKYQDR 180
Db 135 GEYSGWYSVDEBFTTESQLAEVFRDEAGNVTGCIAPSGHEVWVSEESYFLRLSKYQDR 194
Qy 181 LVEFFKAHPFITPDGRLNEMLRNFIEPGLDLAVSRITFTGWVPVPSNPKHVVYWIDA 240
Db 195 LVEFFKAHPFITPDGRLNEMLRNFIEPGLDLAVSRITFTGWVPVPSNPKHVVYWIDA 254
Qy 241 LLNATYALGVAQDEHGFDFWNGTVFHMVKGKILRPHSIYWPILLMMLDVKLPDRLIAH 300
Db 255 LLNATYALGVAQDEHGFDFWNGTVFHMVKGKILRPHSIYWPILLMMLDVKLPDRLIAH 314
Qy 301 GWFVKDGKSKSGNVTYEMLVRYGLDPLRYLNRNLPVSGDGTFTFEDVVGRIYNE 360
Db 315 GWFVKDGKSKSGNVTYEMLVRYGLDPLRYLNRNLPVSGDGTFTFEDVVGRIYNE 374
Qy 361 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTFDFHVLAEVAKSIADFFTHMEAVDYPR 420
Db 375 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTFDFHVLAEVAKSIADFFTHMEAVDYPR 434

Qy 421 ALEAVTTLISRTNKYIDETAPWVLKDDEALRDQALASVMHQAASIRVVVAHLIEPMMETS 480
Db 435 ALEAVTTLISRTNKYIDETAPWVLKDDEALRDQALASVMHQAASIRVVVAHLIEPMMETS 494
Qy 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIPPRLANNEBEIAYIKQMEGNKPAV 540
Db 495 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIPPRLANNEBEIAYIKQMEGNKPAV 554
Qy 541 EXEMNPDEVELKLNKDEIKFDFDKVEIRVAEVEKSVKVEGSKLLOFRLDAGDGEDRQI 600
Db 555 EXEMNPDEVELKLNKDEIKFDFDKVEIRVAEVEKSVKVEGSKLLOFRLDAGDGEDRQI 614
Qy 601 LSGIAKYYPNEGELVKKVQIVANLKPRKMMKKYVSQGMILSAEHDGKLTLLTVDPVAPN 660
Db 615 LSGIAKYYPNEGELVKKVQIVANLKPRKMMKKYVSQGMILSAEHDGKLTLLTVDPVAPN 674
Qy 661 GSVIG 665
Db 675 GSVIG 679

RESULT 4
US-10-282-122A-74415
Sequence 74415, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 74415
LENGTH: 666
TYPE: PRT
ORGANISM: Streptococcus pyogenes
US-10-282-122A-74415

Query Match 82.9%; Score 2883.5; DB 15; Length 666;

Best Local Similarity 81.9%; Pred. No. 1.9e-216;
Matches 546; Conservative 54; Mismatches 64; Indels 3; Gaps 2;
QY 1 MGEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDYFYLGLDEHGKIQ 60
Db 1 MKKPPYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGHEVYFYLGLDEHGKIQ 60
QY 61 ABEAGITPQAYVDGAVGVKELWQLDISYDKFIITDDYHEKVAQVFERLLAQDDIYL 120
Db 61 ABEAGITPQAYVDGAVGVKELWQLDISYDKFIITDDYHEKVAQVFERLLAQDDIYL 120
QY 121 GEYSGHYSVDSDEFFTESQLAHVDEAGNVGTGGIAPSGHEVWVSESYFLRLSKYQDR 180
Db 121 GEYSGHYSVDSDEFFTESQLEKVRDEDOQVIGGIAPSGHEVWVSESYFLRLSKYQDR 180
QY 181 LVEFFKAHPEFITPGRNLNEMRNFIPEGLEDLAVSRITFTWGVVPVSPNPKHVVTWIDA 240
Db 181 LVAFFEREPDFIQDGRNEMVKNFTEPEGLEDLAVSRITFTWGVVPVSPNPKHVVTWIDA 240
QY 241 LNNYATGALGYQAHNFDFKFWNGTVFHVWVGKDILRFHSIYWPILLMMLDVKLPDLIAH 300
Db 241 LNNYATGALGYQAHNFDFKFWNGTVFHVWVGKDILRFHSIYWPILLMMLDVKLPDLIAH 300
QY 301 GWFVMDGKWSKSGNVVYPEMLVEREGLDPLRYLLMRSLPVGSDGTFPTPDYVGRINYE 360
Db 301 GWFVMDGKWSKSGNVVYPEMLVEREGLDPLRYLLMRSLPVGSDGTFPTPDYVGRINYE 360
QY 361 LANDLGNLNRVTSMINKYFDQCI PAYVEGVTEFDFHVLAEVAKSIAADPHTHMEAVDYP 420
Db 361 LANDLGNLNRVTSMINKYFDQCI PAYVEGVTEFDFHVLAEVAKSIAADPHTHMEAVDYP 420
QY 421 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDLQASVMSHWQASIRVVAHLIEPPEMETS 480
Db 421 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDLQASVMSHWQASIRVVAHLIEPPEMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPPLNMBEIIAIKEQMEGNKPAV 540
Db 481 AALMAQLGEPVSDLSLALADFPADVTVVAKGTPIFPPLNMBEIIAIKEQMEGNKPAV 540
QY 541 --BKEWNPDEVELKLNKDEIKFEDFDKVEIRVAEVKEVSKVSGSKLLQPLDAGDGR 598
Db 540 SOEKWNPDEVELKLNKDEIKFEDFDKVEIRVAEVKEVSKVSGSKLLQPLDAGDGR 598
QY 599 QILSGIAKYFNEQELVGGKQIIVANLKPQWKKYISQGMILSAEHGDKLTLVLDVAV 658
Db 600 QILSGIAKYFNEQELVGGKQIIVANLKPQWKKYISQGMILSAEHGDKLTLVLDVAV 658
QY 659 PNGSVTG 665
Db 660 PNGSIIG 666

RESULT 5
US-10-282-122A-72423
Sequence 72423, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Chlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 72423
LENGTH: 669
TYPE: PRT
ORGANISM: Streptococcus mutans
US-10-282-122A-72423

Query Match 82.9%; Score 2883; DB 15; Length 669;
Best Local Similarity 82.4%; Pred. No. 2.1e-216;
Matches 553; Conservative 46; Mismatches 51; Indels 8; Gaps 5;
QY 1 MBEKN-FYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDYFYLGLDEHGKIQ 59
Db 1 MTEKQPYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDYFYLGLDEHGKIQ 60
QY 60 KAEAGITPQAYVDGAVGVKELWQLDISYDKFIITDDYHEKVAQVFERLLAQDDIY 119
Db 61 KSEAGITPQAYVDGAVGVKELWQLDISYDKFIITDDYHEKVAQVFERLLAQDDIY 120
QY 120 LGEYSGHYSVDSDEFFTESQLAHVDEAGNVGTGGIAPSGHEVWVSESYFLRLSKYQDR 179
Db 121 LGEYSGHYSVDSDEFFTESQLEKVRDEDOQVIGGIAPSGHEVWVSESYFLRLSKYQDR 180
QY 180 RLVEFFKAHPEFITPGRNLNEMRNFIPEGLEDLAVSRITFTWGVVPVSPNPKHVVTWID 239
Db 181 RLVDFFHAHPEFITPGRNLNEMRNFIPEGLEDLAVSRITFTWGVVPVSPNPKHVVTWID 240
QY 240 ALLNYATGALGYQAHNFDFKFWNGTVFHVWVGKDILRFHSIYWPILLMMLDVKLPDLIAH 299
Db 241 ALLNYATGALGYQAHNFDFKFWNGTVFHVWVGKDILRFHSIYWPILLMMLDVKLPDLIAH 300
QY 300 HGWFVMDGKWSKSGNVVYPEMLVEREGLDPLRYLLMRSLPVGSDGTFPTPDYVGRINYE 359
Db 301 HGWFVMDGKWSKSGNVVYPEMLVEREGLDPLRYLLMRSLPVGSDGTFPTPDYVGRINYE 360
QY 360 ELANDLGNLNRVTSMINKYFDQCI PAYVEGVTEFDFHVLAEVAKSIAADPHTHMEAVDYP 419
Db 361 ELANDLGNLNRVTSMINKYFDQCI PAYVEGVTEFDFHVLAEVAKSIAADPHTHMEAVDYP 420
QY 420 RALEAVWTLISRTNKYIDETAPWVLDKDEALRDLQASVMSHWQASIRVVAHLIEPPEMETS 479
Db 421 RALEAVWTLISRTNKYIDETAPWVLDKDEALRDLQASVMSHWQASIRVVAHLIEPPEMETS 480
QY 480 SRAVLTOIGLEEVSSLENLSLADFPADVTVVAKGTPIFPPLNMBEIIAIKEQMEGNKPA 539
Db 481 SNAIMEQLGAGQFDLEGLAGLPGVGVIAKGTPIFPPLNMBEIIAIKEQMEGNKPA 539
QY 540 V---BKEWNPDEVELKLNKDEIKFEDFDKVEIRVAEVKEVSKVSGSKLLQPLDAGD 595
Db 540 IQAEEKEWNPDEVELKLNKDEIKFEDFDKVEIRVAEVKEVSKVSGSKLLQPLDAGD 599

QY 596 EDQILSGIAKYPNEQELVKKVQIVANLKPRKMMKKYVSGQMILSASH-DGKLTLLTV 654
DB 600 EDQILSGIAKYPNEQELVKKVQIVANLKPRKMM-KLVSGQMILSASHADGKLTLLTV 658
QY 655 DPAVPSNGSVIG 665
DB 659 DSSVPNGSQIG 669
RESULT 6
US-10-282-122A-57809
; Sequence 10282, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57809
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Enterococcus faecium
US-10-282-122A-57809
Query Match 66.0%; Score 2295; DB 15; Length 669;
Best Local Similarity 65.2%; Pred. No. 2e-170;
Matches 438; Conservative 93; Mismatches 129; Indels 12; Gaps 9;
QY 1 MSEK-NEYITPIYVPSGKLHGSAYTTIACDVLARYKRLMGVDVFLTGLDGHGKIQQ 59
DB 1 MAEKETFYITPIYPSGKLHGSAYTTIACDVLARYKRLMGVDVFLTGLDGHGKIQET 60
QY 60 KAEKAGITPQAYVDGMAVGKELWQLDLSYDKFIRTTDDYHKKVAVQPERLLAODDIY 119
DB 61 KAEELGVKPCQYVDKMAADVKLWKTLDLSYDKFIRTTDDYHKKVAVQKIFDRLLQGGDIY 120
QY 120 LGEYSGHYSVDEEFFTESQLAEVPRDEAGVNTGGIAPSGHEVWSEERSYFRLSKYQD 179

DB 121 LGEYSGHYSVDEEFFTESQLAEVYKDDGNGVIGKAPSGHEVWSEERSYFRLSKYAD 180
QY 180 RLVEFEKAPHPFITPDGLNEMLENFIEPGLDLAVSRITFTWGVVPSPKPHVYVWID 239
DB 181 RLLEYYEHPFITPDGLNEMLENFIEPGLDLAVSRITFTWGVVPSPKPHVYVWID 240
QY 240 ALLNYATAGYAQDEHGNFDFXWNGTVPHVYVWIDILRFHSHYIWPILLMLDVKLPDRLLA 299
DB 241 ALSNYITAGYSEDESLEFEXYFADV-QVVGKEIVRFHTIYWIIMLMALDPLPKKVF 299
QY 300 HGWFVMDKGMKSKGNVYVPEMLVERVYGLDPLAYVYLMNLPVSGDGTFTPEEDVVGRTNY 359
DB 300 HGWLLMDKGMKSKGNVYVPEMLVERVYGLDPLAYVYLMNLPVSGDGTFTPEEDVVGRTNY 359
QY 360 ELANDLGNLNRVTSMINKYFDGQIPAYVEGVTFDFHVAEVAEKSIADFEHMEAVDYP 419
DB 360 DLANDLGNLNRVTSMINKYFDGQIPAYVEGVTFDFHVAEVAEKSIADFEHMEAVDYP 419
QY 420 RALEAVWTLISRTNKYIDETAPWVLDKDBALRDQASVMSHWQASIRVAHLIEPFMMET 479
DB 420 TALAEIWLISRTNKYIDETAPWVLDKDBALRDQASVMSHWQASIRVAHLIEPFMMET 479
QY 480 SRVLTQIGLB-EVSSLNLSLADFPADVTVVAKGTPIFFRLNMBEETIAYIKEQM-EQ-- 535
DB 480 PEKFNQGLDSETNWLGHLFGFPGSGTKVAKGTPIFFRLNMBEETIAYIKEQM-EQ-- 535
QY 536 -NKPAVEKMPDVELKLNKD-EIKFDFDKVIRVAEVAEKSVKVEGSKLQFLRLDAG 593
DB 540 TNEDTV--KWDPEETELVSTKEKQIKFDFVEKVELKVAEINCKVEGADKLQFLRLDAG 597
QY 594 DCEQDQILSGIAKYPNEQELVKKVQIVANLKPRKMMKKYVSGQMILSASH-HDGKLTLL 652
DB 598 DSQDQILSGIAKYPNEQELVKKVQIVANLKPRKMMKKYVSGQMILSASH-HDGKLTLL 652
QY 653 TVDPAVPSNGSVI 664
DB 657 BAPKSMPSNGSEI 668
RESULT 7
US-09-815-242-10663
; Sequence 10663, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 10663
 LENGTH: 677
 TYPE: PRT
 ORGANISM: Enterococcus faecalis
 US-09-815-242-10863

Query Match 65.8%; Score 2288; DB 9; Length 677;
 Best Local Similarity 63.9%; Pred. No. 7.3e-170;
 Matches 428; Conservative 104; Mismatches 130; Indels 8; Gaps 7;

QY 1 MSEK-NFYITPTIYPSGKLHGSAYTTACDVARYKRLMGVDYFYLGLDHEGKIQ 59
 DB 9 MSQKETFYITPTIYPSGKLHGSAYTTACDVARYKRLMGVDYFYLGLDHEGKIEK 68
 QY 60 KAEAGITPOAYVDGMAVGKELWQLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIY 119
 DB 69 KAAELNVTPEYVDKAAADVQKLWKTLDISYDKFIRTTDDYHMAAVQQIFDLRVSGDIY 128
 QY 120 LGEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSSESYFLRLSKYQD 179
 DB 129 LGEYEGWYSVDSDEFFTETQLAEVYRDDGNNVIGGKAPSGHEVELYKEESYFFRMSKYAD 188
 QY 180 RLVEFFKAPHEFTTPDGRNLEMLNFIETPCGLDLAVSRITFTTGWVPSPNPKHVYVWID 239
 DB 189 RLVOYEEHPEFTQPSERKNEMLNFIETPCGLDLAVSRITFTTGWVPSPNPKHVYVWID 248
 QY 240 ALLNYATGALYAQDEHGNFDFKFWNGTIVFHMVGKDLIRFHSIYWPILLMMLDVKLPDLRIA 299
 DB 249 ALSNYITALGYSEDSDSLFQKYWPANV-QMVGKEIVRFHTIYWPIMLMALDPLPKKVF 307
 QY 300 HGWFMKDGKMSKSGNVVYPEMLVERYGIDPLRYVLMENLPVSGDGTTPEDYGRINY 359
 DB 308 HGWLLMKDGMKSKSGNVVYPEMLVERYGIDPLRYVLMENLPVSGDGTTPEDYGRINY 367
 QY 360 ELANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVLAEVAEKSIADPHTMEAVDYP 419
 DB 368 DLANDLGNLNRVTSMINKYCDGKVPAYASKVTPFDSLSSTTAANVIGKYHEAMEKVEFN 427
 QY 420 RALEAVWTLISRTNKYIDETAPWVLDKDEALDOLASVMSHQASIRVVAHLIEPFMMET 479
 DB 428 TATAEITWLVSRANKYIDETAPWVLAKEEKEKNELESVMHIAESLRIVAILLQPVMTET 487
 QY 480 SRAVLITQIGLE-EVSSLENLSLADFPADVTVVAKGTPIFPRLNMBEEIAYIKEQMEGNKP 538
 DB 488 PGKIFQGLDPTNMENIHFGEFTDVTNISKGTPIFPLRIETEVTYIQKMSQSES 547
 QY 539 AVEKE--WNPDEVELKIND-EIKFEDFDKVEIRVAEVEKSVGSGKLLQFRLDAGDG 595
 DB 548 ATBEDIKWNPDEETLVSTREKQIKYDDFDKVELKVAEVIDCKKVKGADKLLQFRLDAGDE 607
 QY 596 EDQIISLGIATKYPNQEVLGVKQVQIVANLKPRLMKKVKVVSQGMILSABH-DGKLTLLTV 654
 DB 608 NHRQIISLGAIEFPDPAALIGKXWIVANLKPRL-MRQIISQGMILSABSPGKLIQIVEA 666
 QY 655 DPVAPNGSVI 664
 DB 667 PKEMPNGAGI 676

RESULT 8

US-10-282-122A-42475
 Sequence 42475, Application US/10282122A
 Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Hasebeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA.034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-07-27
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-16
 PRIOR APPLICATION NUMBER: 60/269,308
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 42475
 LENGTH: 677
 TYPE: PRT
 ORGANISM: Enterococcus faecalis
 US-10-282-122A-42475

Query Match 65.8%; Score 2288; DB 15; Length 677;
 Best Local Similarity 63.9%; Pred. No. 7.3e-170;
 Matches 428; Conservative 104; Mismatches 130; Indels 8; Gaps 7;

QY 1 MSEK-NFYITPTIYPSGKLHGSAYTTACDVARYKRLMGVDYFYLGLDHEGKIQ 59
 DB 9 MSQKETFYITPTIYPSGKLHGSAYTTACDVARYKRLMGVDYFYLGLDHEGKIEK 68
 QY 60 KAEAGITPOAYVDGMAVGKELWQLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIY 119
 DB 69 KAAELNVTPEYVDKAAADVQKLWKTLDISYDKFIRTTDDYHMAAVQQIFDLRVSGDIY 128
 QY 120 LGEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSSESYFLRLSKYQD 179
 DB 129 LGEYEGWYSVDSDEFFTETQLAEVYRDDGNNVIGGKAPSGHEVELYKEESYFFRMSKYAD 188
 QY 180 RLVEFFKAPHEFTTPDGRNLEMLNFIETPCGLDLAVSRITFTTGWVPSPNPKHVYVWID 239
 DB 189 RLVOYEEHPEFTQPSERKNEMLNFIETPCGLDLAVSRITFTTGWVPSPNPKHVYVWID 248
 QY 240 ALLNYATGALYAQDEHGNFDFKFWNGTIVFHMVGKDLIRFHSIYWPILLMMLDVKLPDLRIA 299
 DB 249 ALSNYITALGYSEDSDSLFQKYWPANV-QMVGKEIVRFHTIYWPIMLMALDPLPKKVF 307
 QY 300 HGWFMKDGKMSKSGNVVYPEMLVERYGIDPLRYVLMENLPVSGDGTTPEDYGRINY 359
 DB 308 HGWLLMKDGMKSKSGNVVYPEMLVERYGIDPLRYVLMENLPVSGDGTTPEDYGRINY 367
 QY 360 ELANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVLAEVAEKSIADPHTMEAVDYP 419
 DB 368 DLANDLGNLNRVTSMINKYCDGKVPAYASKVTPFDSLSSTTAANVIGKYHEAMEKVEFN 427
 QY 420 RALEAVWTLISRTNKYIDETAPWVLDKDEALDOLASVMSHQASIRVVAHLIEPFMMET 479
 DB 428 TATAEITWLVSRANKYIDETAPWVLAKEEKEKNELESVMHIAESLRIVAILLQPVMTET 487
 QY 480 SRAVLITQIGLE-EVSSLENLSLADFPADVTVVAKGTPIFPRLNMBEEIAYIKEQMEGNKP 538

```
Db 488 PGKIFQGLDPTMTNMENIHFGFTDVTVTSGTPIFPRLIEITEVYIYOKMSQSES 547
Qy 539 AVEKE--WNPDEVELKLNKD-EIKFEDFDKVEIRVAEVEKSVKSGSKLLQFRLDAGDG 595
Db 548 ATEEDIKNPEETTLVSTKEKQIKYDDFDKVELKVAEVIDCKVKGADKLLQFRLDAGDE 607
Qy 596 EDRQILSGTAKYYPNEQELVGKQVQIVANLKPRKMMKKVVSQGMILSAEH-DGKLTLLTV 554
Db 608 NHRQILSGTAEAFYDPAALIGKVKVIVANLKPRK-MRQGISQGMILSAESPEGKIQIVEA 666
Qy 655 DPAVNGSVI 664
Db 667 PKEMPNGAGI 676

RESULT 9
US-09-815-242-4933
; Sequence 4933, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 4933
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-4933

Query Match 65.7%; Score 2284.5; DB 9; Length 666;
Best Local Similarity 63.7%; Pred. No. 1.3e-169;
Matches 425; Conservative 105; Mismatches 130; Indels 7; Gaps 6;

Qy 3 EKFPYITPIYPSGKLHIGSAYTTIACDVLARYKRLMGYDVFYLTGLDGHGKIOQKAE 62
Db 2 KETFYITPIYPSGKLHIGSNYTTIACDWARVKRLMGDFVYLTGLDGHGKIEKAA 61
Qy 63 EAGITPAYDGMAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAODDIYLGE 122
Db 62 ELNVTPEYVDKMAADVQLWKLTLDISYDKFIRTTDDYHMAAVQOIFDLRVEQGDYILGE 121
Qy 123 YSGWYSVDEBFFTESQALAEVFRBAGNVGGTAPSGHEVWVSESYFRLSKYODRLV 182
Db 122 YEGWYSVDEBFFETQALAEYRDEGNVGGKAPSGHEVVKESYFFRMSKYADRLV 181
Qy 183 EFFKAHPEFITPDGRLEMLRNFTPEGLIEDLAVSRRTTFTWGVVPSPKPHVYVWIDALL 242
```

```
Db 182 QYSEHPEFTQPSERKKNLNNFKPGLEDLAVSRRTTFSNGIPKNDPKHVVYVIDALS 241
Qy 243 NYATALGYAOQEHGNFKFWNGTIVFHMVQXDIILRPHSIYWPILLMLLDVLPDLIAHG 302
Db 242 NYITALGYSSDDSLFQKYNPANY-QMVGEIVAFHTIYWFIMLMALDLPKPKVFGHW 300
Qy 303 FVMKDGKMSKSGNVVPEMLVERYGDLPLRYVLMRNLPGVSGDGTFTPTEDYVGRINVELA 362
Db 301 LLMKDGKMSKSGNVVPEMLVERYGDLPLRYVLMRNLPGVSGDGTFTPTEDYVGRINVELA 360
Qy 363 NDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDYPRAL 422
Db 361 NDLGNLNRVTSMINKYCDGKVPAYASKVTPDFSELSTTAANVIGKYHEAMEKMEFTAI 420
Qy 423 EAVVTLISRTNKYIDETAPVWLDKDEALRDOLASVMSHWQASIRVAHLTEPPFMETSRA 482
Db 421 AEIWTLSVRANKYIDETAPVWLAKEEKEEKEEKEEKEEKEEKEEKEEKEEKEEKEEKEE 480
Qy 483 VLTQIGLE-EVSSLENLSLADFPADVTIVAKGTPIFPRLANMEEETIYIKMEQGNKPAVE 541
Db 481 IFEQLGLDPTMTNMENIHFGFTDVTVTSGTPIFPRLIEITEVYIYOKMSQSESATE 540
Qy 542 KE--WNPDEVELKLNKD-EIKFEDFDKVEIRVAEVEKSVKSGSKLLQFRLDAGDGEDR 598
Db 541 EDIKWNPEETTLVSTKEKQIKYDDFDKVELKVAEVIDCKVKGADKLLQFRLDAGDENHR 600
Qy 599 OILSGIAKYYPNEQELVGKQVQIVANLKPRKMMKKVVSQGMILSAEH-DGKLTLLTVDDPA 657
Db 601 QILSGIAEYDPAALIGKVKVIVANLKPRK-MRQGISQGMILSAESPEGKIQIVEAPKE 659
Qy 658 VPNGSVI 664
Db 660 MPNGAGI 666

RESULT 10
US-10-282-122A-60385
; Sequence 60385, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
```

```

; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60385
; LENGTH: 664
; TYPE: PRF
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60385

Query Match
  60.7%; Score 2113; DB 15; Length 664;
Best Local Similarity
  60.7%; Pred. No. 3.4e-156;
Matches 407; Conservative 101; Mismatches 144; Indels 18; Gaps 8;

Qy 3 EKN-FYITPIYPSGKLHIGSAYTTIACDVLARYKRLMGYDVLYLTGLDEHGQKIQKA 61
Db 4 EKNFTVITPIYPSGKAHIGAYTTVAGDARARYKLGVDVLYLTGDEHGQKIQKA 63
Qy 62 BEAGITPQAYVDMGVKELQQLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 121
Db 64 KERGISEQVVDIARGFQELWKKLEISNTDFIRTTQDRHKTSEKIFQELQEGDIYL 123
Qy 122 EYSGWVSDEBFFTESQLAEVFRDEAGNVGTGIAPSGHEVWSEESYFLRLSKYQDRL 181
Db 124 EYEGWVSDEEYFTQLEEVYKDENGKVIQKAPSGNEVELVKEESYFFKSKYADRL 183
Qy 182 VEFFKAHPFITPDGRLNEMRNFIPEGLEDLAVGRTTFTWGVPSNPKHVVYWDAL 241
Db 184 VEYNHSPFIFLPSKRNEMINNFIPGLEDLAVGRTTFTWGVPSNPKHVVYWDAL 243
Qy 242 LNAVATAGVAQDEHGFDFKNGTFVHMVGKDLAPHSIYWPILLMMDLVKLPDLIAHG 301
Db 244 SNYITAGVNTNDTKFQKYPADV-QVKGKIVRPHIYWPILLMMDLVKLPDLIAHG 302
Qy 302 WFMKDGKMSKSGNVVYPEMLVRYGLDPLRYILMRNLVPGSDGTFPTPDYVGRINYL 361
Db 303 WILKDGKMSKSGNVVDPVMLIDRYGLDALRYLLREVPPGSDGLFTPEDVDVAVNFDL 362
Qy 362 ANDLGNLNRVTSMINKYFDGQIPAYVEGTFEDHVAEVAEKSIADFTHTMEAVDIPRA 421
Db 363 ANDLGNLNRVTAMINKYFDGEIPAYQGNVTFDQILVDKNNVKEVEGSMHMQFSVA 422
Qy 422 LEAVMTLISRTNKYIDETAPWVLDDKDEALRDOLASVMSHWQASIRVAHLIEPFWMTSR 481
Db 423 LNLQWLSLISRTNKYIDETAPWVLDDKDEALRDOLASVMSHWQASIRVAHLIEPFWMTSR 482
Qy 482 AVTQGLSEVSLENL-----SLADPADVTVVAKGTPIFPRLNMEETIAYIKOME 534
Db 483 EIPQLQGLQES---ENLKWDSIYGYGEIPAGTTVVKGTPFPRLDAEVEVYIQDEM 538
Qy 535 GKNPAVEKWNPDVEELKLNKDEIKFEDDKVEIRVAEVEKYSKVEGDKLLQFLDAGD 594
Db 539 GSAPAPAES--VAEVB-ALETQIGIEDFDKIDLRVAEVEKVDKVKKADKLCFOLDGE 595
Qy 595 GEDRQILSGIAYYPNEQELVGKQVQIVANLPRKMKKYVSQGMILSAEHDGKLTIV 654
Db 596 GKLRQVLSGIABFYQPE-ELIGKKIIVSNLKPVK-LRGLMSEGMLSGDKGLSVIEA 653
Qy 655 DPAVNGSVI 664
Db 654 SSALPNGAKV 663

```

RESULT 11

```

US-10-282-122A-71524
; Sequence 71524, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguen
; APPLICANT: Zamudio, Carlos

```

```

; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71524
; LENGTH: 659
; TYPE: PRF
; ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-71524

```

```

Query Match
  56.0%; Score 1949; DB 15; Length 659;
Best Local Similarity
  57.3%; Pred. No. 2.3e-143;
Matches 383; Conservative 103; Mismatches 168; Indels 14; Gaps 9;

Qy 1 MSEKNFVITPIYPSGKLHIGSAYTTIACDVLARYKRLMGYDVLYLTGLDEHGQKIQK 60
Db 1 MAKETFTVITPIYPSGNLHIGAYSTVAGDVLSRYKMGQYDVLYLTGDEHGQKIQK 60
Qy 61 ABEAGITPQAYVDMGVKELQQLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 120
Db 61 AORAGKTELEYLDEMISGKLSKLEISNDDFIRTTEDRHQVKEVFERLLKQGDIVL 120
Qy 121 GEYSGWVSDEBFFTESQLAEVFRDEAGNVGTGIAP-SGHEVEWVSSESYFLRLSKYQD 179
Db 121 GEYEGWVSDEYTYTYSQVLDPIY-ENKIVGKSPDSGHEVELVKEESYFFNINKYTD 179
Qy 180 RLVEFFKAHPFITPDGRLNEMRNFIPEGLEDLAVSRTTFTWGVPSNPKHVVYWD 239
Db 180 RLLEFFDANPDFIOPPSRKNEMINNFIPGLEDLAVSRTSFDWGVVPSNPKHVVYWD 239
Qy 240 ALLNATAGVAQDEHGFDFKNGTFVHMVGKDLRPHYSIYWPILLMMDLVKLPDLIA 299
Db 240 ALVNYISALGYLSDDELQKYPADV-HLMAKEIVRPHYSIYWPILLMMDLVKLPDLIA 298
Qy 300 HGMVFMKDGKMSKSGNVVYPEMLVRYGLDPLRYILMRNLVPGSDGTFPTPDYVGRIN 359
Db 299 HGMILMDKDGKMSKSGNVVDPVMLIDRYGLDALRYLLREVPPGSDGLFTPEDVDV 358
Qy 360 ELANDLGNLNRVTSMINKYFDGQIPAYVEGTFEDHVAEVAEKSIADFTHTMEAVDIP 419
Db 360 ELANDLGNLNRVTSMINKYFDGQIPAYVEGTFEDHVAEVAEKSIADFTHTMEAVDIP 419

```

359 DLANDLGNLVRTISWINKYFOGELPAYEGPKHELDMEALAHETVVKHFNESMESPQS 418
420 RALEAVVTLISRTNKYIDTAPWLDKDBALDQALASVNSHWQASIRVVAHUIEPFMKMT 479
419 VALSTVWKEISRTNKYIDSTTEPWLAQDSQKMDLGNVAHLVENIRFAAVLLRPFLTHA 478
480 SRAVLTOIGLEVSLENLSLADFPRA--DVTVAAGTDFPRLNMBEIEIYKEQMBGN 536
479 PKEIFQLNINPELPELSLSOYGALKQPIWVTEKPTDFPRLDTEAEIYKESMOPP 538
537 KPAVEKEMPDEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVEGSDKLLQFRLDAGDGE 596
539 K-----SEESKDEVE-RPSKAQIDIKDFDKVEIKAAITIDAENVKPSDKLLKQIDLG-LE 592
597 DROILSGIAKYPNEBELGKKVQIVANLKPRKMKKYVSQGMILSAEHGDKLTLLTWDVP 656
593 QRCIVSGIAKFY-RPEDIIIGKVAVVTLNPKAKMGQ-KSEGMILSAEKDGVTLTIVSUPS 650
657 AVPNGSVI 664
651 AIPNGAVI 658

RESULT 12
US-10-282-122A-46123
; Sequence 46123, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-16
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; -SOFTWARE: Patent in version 3.1
; SEQ ID NO 46123
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-282-122A-46123

Query Match 55.6%; Score 1935.5; DB 15; Length 669;
Best Local Similarity 57.0%; Pred. No. 2.7e-142;
Matches 383; Conservative 107; Mismatches 161; Indels 21; Gaps 12;
QY 1 MSEKX--FYITPTIYPSGKLHIGSATTTACDVLAARYKELMGYDVLYLGLDEHGOKIQ 58
10 MTEENKSFYITPTIYPSGKLHIGSATTTACDVLAARYKELMGYDVLYLGLDEHGOKIQ 69
QY 59 QKAEAGITTPQAYVDGMAGVGMKELQQLDISYDKFIRTTDDYHEKVAQVAFERLLAQDDI 118
70 KKAELNVTPOAYVDNIVAGIKELWEXMDISYDDFTTTEDRHKDVVEKIFKQLVDQGD 129
119 YLGEYSGMYSVDEEFFTESQALAEVFRDEAGNTVGGAP-SGHEVEWVWSEESFLRLSKY 177
130 YLDEVEGWYSVDTEFTYEHQVDPIM-EGDKVVGKSPDSDGHDVLELRESEYFFRMGY 188
178 QDLVEFKAHPEITPDGRNLNEMLRNFIIPGLDLAVSRFTTGTWGVPSNPKHVVYVW 237
189 VDRLLKFYEDNPHFIQEPESRKNEMINNFIPGLDLAVSRFTTGTWGVPSNPKHVVYVW 248
238 IDALLNATAGYAQDBEHGDFKFNWGTVPFHMVGKDLIRPHSIIYWPILLMMKLDVKLPDL 297
249 VDALSNTITAGYGTENEEMYYKFWPADV-HLVGKEIVRPHTIYWPILLMALDPLPKKV 307
298 IAHGFWNKGKMSKSGNVVYPEMLVERIGLDPIRYLMLNLPVSGDGTETPDYVGR 357
308 FAHGWLKMDGMSKSGNVVDPVILIDRYGLDALRYLLREVDFGSDGVTPGDFVER 367
358 NYELANDLGNLNRVTSMINKYFDGQIPAYVEGVTDFDHVLAEVAEKSIADPFHTMEAVD 417
368 NFDLANDLGNLNRVTAMIDKYFNGEIPAFKANVTEDETLVTFKDTLKKVEEAMENNE 427
418 YPRALAVVTLISRTNKYIDETAPWLDKDBALDQALASVNSHWQASIRVVAHUIEPFM 477
428 FVALSSIQVLSVSRNKYIDETQPVWLAQDDDEKLAASVNSHWQASIRVVAHUIEPFM 487
478 ETSRAVLTOIGL--REVSSLENLS-LADFPADVTVWAKGTPIFPRLNMBEIEIYKEQME 534
488 VAPSKQFAQLGTLTDEAHTSWGSLSTIGCIPAG-TKVEKGQIPFRLDNDVEVAIKQMK 546
535 GNKPAVE--KEWNPDEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVEGSKLQFLRDA 592
547 ASAPKVEEKKEEP-----KAEITITDFFKVELRVAEVLSEAPVKKADKLLKIQJDL 599
593 GDGEDROILSGIAKYPNEBELGKKVQIVANLKPRKMKKYVSQGMILSAEHGDKLTLL 652
600 GT-EXRQVVGIAKFY-SPEDLKGGKVICVTLNPKVK-LRGELSGQMLAGEENGVLSLA 656
653 TVDPAVPNGSVI 664
657 SIDQNPNGTKI 668
RESULT 13
US-09-815-242-12135
; Sequence 12135, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21


```

; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12135
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-12135

Query Match      55.6%; Score 1934; DB 9; Length 657;
Best Local Similarity 56.3%; Pred. No. 3.4e-142;
Matches 376; Conservative 111; Mismatches 165; Indels 16; Gaps 10;

QY      1 MBEKNFYITPIYPSPGKHLIGSAVTIACDVLARYKRLMGYDVFLYGLDDEHGQKIQOK 60
DB      1 MAKEETIITPIYPSPGNLHIGHAYSTVAGDVIARYKRMQGYDVFLYGLTDEHGQKIQEK 60

QY      61 ABEAGITPQAYVDGMAGVKELWQLLDISYDKPIRTDDYHEKVAQVFERLLAODDIYL 120
DB      61 AOKAGKTEIYLDDEMTAGIKQLWAKLEISNDDFIRTEERHKHVVEQVFERLLKQGDYIL 120

QY      121 GEYSGWYSVDEEFFTESQLAEVFRDEAGNVGGIAP-SGHEVWVSSEYFLRLSKYOD 179
DB      121 GEYEGWYSVDETYTYESQLVDP-QVENGKIIGKSPSGSHEVELVKESYFFNISKYTD 179

QY      180 RLVEFFKAHPFITPDGRINEMLRNFIPEGLDLAVSRFTETWGVPSNPKNHVVYVWID 239
DB      180 RLLEFFDQNPDTIQPPSRKNEMINFIKGLADLAVSRFTSNWGVHVPSPNPKHVVYVWID 239

QY      240 ALLNYATAGYAQDEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMMLDVKLPDLRIA 299
DB      240 ALVNYISALGYLSDESLEFNKYWPADI-HLMAKEIVRFHSIITWPIILLMALDPLPKKVFA 298

QY      300 HGVFVNMKGKMSKGNVYYPPEMLVRYGLDPLRYLLMRNLPGVSDGTFTPEDYVGRINY 359
DB      299 HGWILMKDGKMSKGNVVDNILLIDRYGLDATRYLLMRNLPFGSDGVFTPEAFVETNF 358

QY      360 ELANDIGNLNETVSMINKYFGQIPAYVEGTEDPHVLAEVAEKSIADFFTHMEANDYP 419
DB      359 DLANDIGNLVNRTISMINKYFGGELPFAVQGPPLHDEMEAMALTEVKSYSMESLQFS 418

QY      420 RALEAVWTLSIRTNKYIDETAPWLKDDEALDQLASVMSHWQASIRVVVAHLIRPFWMET 479
DB      419 VALSIWKFIISRTNKYIDETTPVLAKODSQKMLGNVMAHLVENIRVAALLAPFLTHA 478

QY      480 SRAVITQIGLEEVSSLENIUSLADF--PADVTVVAKGTPPIPRLNMBEIRIAYIKQEBGN 536
DB      479 PKEIPEQLINNPFQMEFSLAQYGLVTEPIWVTQPKPIPRLDSEABEIRIAYIKESMQ-- 536

QY      537 KPAVEKMNPDVBELKLNKDEIKFDEFDKVEIRVAEVEKSVESGSKLLOFRLDAGDGE 596
DB      537 PRATEEE--KEIIP--SKPQIDIKDFKVEIKAAITIDAEHVKSCKLLKIQVDL-DSE 590

QY      597 DRQILSGIAKYYPNEQELVGKVKQIVANLKPRKMKTKYVSGMILASBHGKLTLLTYDP 656
DB      591 QRQIVSGIAKPY-TPDDIIGKKVAVVTLNPKAKLMGQ-KSEGMILASBKDGVLTLVSLPS 648

QY      657 AVPNGSVI 664
DB      649 ALPNGAVI 656

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 43933
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-10-282-122A-43933

Query Match      55.5%; Score 1931; DB 15; Length 657;
Best Local Similarity 56.3%; Pred. No. 5.8e-142;
Matches 377; Conservative 110; Mismatches 163; Indels 20; Gaps 11;

QY      1 MBEKNFYITPIYPSPGKHLIGSAVTIACDVLARYKRLMGYDVFLYGLDDEHGQKIQOK 60
DB      1 MAKEETIITPIYPSPGNLHIGHAYSTVAGDVIARYKRMQGYDVFLYGLTDEHGQKIQEK 60

QY      61 ABEAGITPQAYVDGMAGVKELWQLLDISYDKPIRTDDYHEKVAQVFERLLAODDIYL 120
DB      61 AOKAGKTEIYLDDEMTAGIKQLWAKLEISNDDFIRTEERHKHVVEQVFERLLKQGDYIL 120

QY      121 GEYSGWYSVDEEFFTESQLAEVFRDEAGNVGGIAP-SGHEVWVSSEYFLRLSKYOD 179
DB      121 GEYEGWYSVDETYTYESQLVDP-QVENGKIIGKSPSGSHEVELVKESYFFNISKYTD 179

QY      180 RLVEFFKAHPFITPDGRINEMLRNFIPEGLDLAVSRFTETWGVPSNPKNHVVYVWID 239
DB      180 RLLEFFDQNPDTIQPPSRKNEMINFIKGLADLAVSRFTSNWGVHVPSPNPKHVVYVWID 239

QY      240 ALLNYATAGYAQDEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMMLDVKLPDLRIA 299
DB      240 ALVNYISALGYLSDESLEFNKYWPADI-HLMAKEIVRFHSIITWPIILLMALDPLPKKVFA 298
```

QY 300 HGWFMKDGKMSKGNVYVPEMLVERVGLDPLRYVLMRNLPVSGDGTFTPDYGRINY 359
Db 299 HGWILMKDGKMSKGNVVDNLIIDRYGLDTRVYLMRELPGSDGVFTPEAFVERTNF 358
QY 360 ELANDLGNLNRVTSMINKYFDGOIPAYVEGVTEFDHVLAEVAEKSIADFTHMEAVDYP 419
Db 359 DLANDLGNLVNRTISWNKYFDGELPAYQGPLHELDEEMEALETVKSYTESMESLQFS 418
QY 420 RALEAVWTLISRTNKYIDETAPVWLDKDEALRDLQASVMSHWQASIRVVAHLIEPMMET 479
Db 419 VALSTVWKFIKRTNKYIDETTPVWLAKDSDQMDLGNVMAHLVENIRYAVALRPLFLTHA 478
QY 480 SRAVLTIQIGLE-----EVSSLENLSLADFPADVTVVAKGTPIPPRLNMEEBIAYIKEOME 534
Db 479 PKEIFEQLNINNPQFMESLSLEQYGVN--ESIMVTGQPKDIFPRLDSEABIAYIKESMQ 536
QY 535 GNKPAVEKEWNPDEVELKLNKDEIKFEDPKVEIRVAEVEKSVKVGSDKLLQFRLDAGD 594
Db 537 --PPATEEE--KEEIP---SKPQIDIKDFDKVEIKAATIIDAEHVKKSKLLKIQVDL-D 588
QY 595 GEDROIISGLIAKYYPNEQELVGHKKVQIVANLKPROMKKYVSGMILSAEHDGKLTLLTV 654
Db 589 SEQROIIVSGIAKFI--TPDDIIGKVVAVTNLKPALMGQ--KSEGMILSAEKDGVLTLLVSL 646
QY 655 DPAVNGSVI 664
Db 647 PSAIPNGAVI 656

RESULT 15

US-09-815-242-13013
; Sequence 13013, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/131,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13013
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13013

Query Match 55.5%; Score 1930; DB 9; Length 657;
Best Local Similarity 56.1%; Pred. No. 6.9e-142;
Matches 376; Conservative 108; Mismatches 166; Indels 20; Gaps 9;

QY 1 MSEKNFYITTPYPPSGKLHIGSAYTTIACDVLARYKRLMGDYVYLTGLDEHQKIQOK 60
Db 1 MAKTFITTPYPPSGNLHIGHAYSIVAGDVARYKEMQGYVRYLTGTDEHQKIQOK 60
QY 61 AEEAGITPOAYVDMAYGVXELMOLLDSYDKFTRTDDHEKVAQVAFERLLAODIYL 120
Db 61 AOKAGKTEIEYLDDEMIAGIKQMAKLEISNDDFTIRTEERHKHYVEQVFERLLAQGIYL 120
QY 121 GEYSGWYSVDEBEFTESQLAEVFRDEAGNVGTGIAP--SGHEVEWVSEESYFRLSKYQD 179
Db 121 GEYSGWYSVPDETYYTSSQLVDP--QYENGKLIIGKSPDSGHEVELVKEESYFFNISKYTD 179
QY 180 RLVEFFFKAHPEFTITPDGRLEMLENFIEPGLLEDLAVSRTTFTMGVPPVPSNPKHVVYVWD 239
Db 180 RLLEFYDQNPDIQPPSKQENMINNFIKPGLADLAVSRTSFNMGVHPVPSNPKHVVYVWD 239
QY 240 ALLNYATRALGVAQAEHGNDFKFNNGTVFHMVKDILSFHSIYWPILLMLDVLKLPDLRIA 299
Db 240 ALVNYISALGYSLSDESLEFNKYWPADI--HLMAKSIIVRFHSIIFWILLMALDLPKPKVFA 298
QY 300 HGWFMKDGKMSKGNVYVPEMLVERVGLDPLRYVLMRNLPVSGDGTFTPDYGRINY 359
Db 299 HGWILMKDGKMSKGNVVDNLIIDRYGLDTRVYLMRELPGSDGVFTPEAFVERTNF 358
QY 360 ELANDLGNLNRVTSMINKYFDGOIPAYVEGVTEFDHVLAEVAEKSIADFTHMEAVDYP 419
Db 359 DLANDLGNLVNRTISWNKYFDGELPAYQGPLHELDEEMEALETVKSYTESMESLQFS 418
QY 420 RALEAVWTLISRTNKYIDETAPVWLDKDEALRDLQASVMSHWQASIRVVAHLIEPMMET 479
Db 419 VALSTVWKFIKRTNKYIDETTPVWLAKDSDQMDLGNVMAHLVENIRYAVALRPLFLTHA 478
QY 480 SRAVLTIQIGLE-----EVSSLENLSLADFPADVTVVAKGTPIPPRLNMEEBIAYIKEOME 534
Db 479 PKEIFEQLNINNPQFMESLSLEQYGVN--ESIMVTGQPKDIFPRLDSEABIAYIKESMQ 536
QY 535 GNKPAVEKEWNPDEVELKLNKDEIKFEDPKVEIRVAEVEKSVKVGSDKLLQFRLDAGD 594
Db 537 PPATKEEKEEIP-----SKPQIDIKDFDKVEIKAATIIDAEHVKKSKLLKIQVDL-D 588
QY 595 GEDROIISGLIAKYYPNEQELVGHKKVQIVANLKPROMKKYVSGMILSAEHDGKLTLLTV 654
Db 589 SEQROIIVSGIAKFI--TPDDIIGKVVAVTNLKPALMGQ--KSEGMILSAEKDGVLTLLVSL 646
QY 655 DPAVNGSVI 664
Db 647 PSAIPNGAVI 656

Search completed: December 17, 2004, 22:11:32
Job time : 151 secs

This Page Blank (uspto)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	3446	99.1	665	2	D95091	methionyl-tRNA syn
2	3446	99.1	679	2	H97958	methionine-tRNA li
3	2392.5	68.8	662	2	H86723	methionine-tRNA li
4	2113	60.7	664	2	AB3097	methionyl-tRNA syn
5	2107	60.6	664	2	A11459	methionyl-tRNA syn
6	1363.5	56.4	684	2	S60667	methionine-tRNA li
7	1953.5	56.2	660	2	E83656	methionyl-tRNA syn
8	1931	55.5	657	2	C89815	methionyl-tRNA syn
9	1919	55.2	649	2	S16682	methionine-tRNA li
10	1640.5	47.2	644	2	A97268	methionyl-tRNA syn
11	1359.5	39.1	629	2	E73297	methionine-tRNA li
12	1189.5	34.2	628	2	F81356	methionine-tRNA li
13	1153.5	33.2	521	2	C89222	methionyl-tRNA syn
14	1152.5	33.1	656	2	H71867	methionine-tRNA li
15	1148.5	33.0	650	2	A64572	methionine-tRNA li
16	1140	32.8	497	2	G70408	methionine-tRNA li
17	1131	32.5	508	2	F97830	methionine-tRNA li
18	1129	32.5	616	1	SYTMT	methionine-tRNA li
19	1068.5	30.7	516	2	F97541	methionyl-tRNA syn
20	1068.5	30.7	516	2	AH3760	methionyl-tRNA syn
21	1062	30.5	508	2	F71674	methionine-tRNA li
22	1058	30.4	544	2	AE3375	methionine-tRNA li
23	1047.5	30.1	530	2	A11835	methionyl-tRNA syn
24	1040	29.9	537	2	F89338	probable methionyl
25	1037.5	29.8	532	2	S75525	methionine-tRNA li
26	1037	29.8	519	2	B70603	probable methionyl
27	1018	29.3	681	2	C75395	methionyl-tRNA syn
28	974	28.0	509	2	B90539	hypothetical prote
29	904	26.0	616	2	T47679	methionine-tRNA li

```

Db 361 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEQSIADPHTEAVDYPR 420
QY 421 ALBAVWTLISRTNKYIDETAPWVLDKDEALRDOLASVMSHMOASIRVVAHLIEPFMMETS 480
Db 421 ALBAVWTLISRTNKYIDETAPWVLDKDEALRDOLASVMSHMOASIRVVAHLIEPFMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
Db 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
QY 541 EKWNPDVELKLNKDBIKFEDFKVEIRVAEVEKVSQVSGDKLLQFLDAGDGEDRQI 600
Db 541 EKWNPDVELKLNKDBIKFEDFKVEIRVAEVEKVSQVSGDKLLQFLDAGDGEDRQI 600
QY 601 LSGIAKYYPNEQELVGKKYQIVANLKPRKMKKKYVSQGMILSAEHDGKLTLLTVDPAPVN 660
Db 601 LSGIAKYYPNEQELVGKKYQIVANLKPRKMKKKYVSQGMILSAEHDGKLTLLTVDPAPVN 660
QY 661 GSVIG 665
Db 661 GSVIG 665

RESULT 2
H97958
methionine-tRNA ligase (EC 6.1.1.10) [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 23-Dec-2002
C;Accession: H97958
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burchett, S.; DeHoff, B.S.; E
y, P.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAnren, S.; M
y, P.; Sun, P.M.; Winkler, M.B.
J. Bacteriol. 183, 5703-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: H97958
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-679 <KOR>
A;Cross-references: GB:AE007317; PIDN:AAK99500.1; PID:g15458285; GSPDB:GN00174
C;Genetics:
A;Gene: metS
C;Superfamily: methionyl-tRNA synthetase, dimer-forming
C;Keywords: ligase

Query Match 99.1%; Score 3446; DB 2; Length 679;
Best Local Similarity 99.2%; Pred. No. 2.4e-198;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 MSEKNFYITPIYPSGKLIHGSAYTTIACDVLARYKRLMGVDVFLYGLDGHGKIQOK 60
Db 15 MSEKNFYITPIYPSGKLIHGSAYTTIACDVLARYKRLMGVDVFLYGLDGHGKIQOK 74
QY 61 ABEAGITPOAYVDGNAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAODDIYL 120
Db 75 ABEAGITPOAYVDGNAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAODDIYL 134
QY 121 GYSYSGHYSVDSDEFFTESQAEVFEDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDR 180
Db 135 GYSYSGHYSVDSDEFFTESQAEVFEDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDR 194
QY 181 LVEFFKAHPEFTTPDGRNLEMLNFIPEGLDLAVSRITFTTGVGPVPSNPKHVVVWIDA 240
Db 195 LVEFFKAHPEFTTPDGRNLEMLNFIPEGLDLAVSRITFTTGVGPVPSNPKHVVVWIDA 254
QY 241 LLNATYALGYAODEHGNFDFKFWNGTVFMVKGDKILRFHSIYWPILLMMLDKVLPDLIAH 300
Db 255 LLNATYALGYAODEHGNFDFKFWNGTVFMVKGDKILRFHSIYWPILLMMLDKVLPDLIAH 314
QY 301 GWFVMDGKMSKGNVYVPEMLVERYGLDPLRYLLMRNLPGVSDGTFTPDYVGRINYE 360

```

```

Db 315 GWFVMDGKMSKGNVYVPEMLVERYGLDPLRYLLMRNLPGVSDGTFTPDYVGRINYE 374
QY 361 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEQSIADPHTEAVDYPR 420
Db 375 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEQSIADPHTEAVDYPR 434
QY 421 ALBAVWTLISRTNKYIDETAPWVLDKDEALRDOLASVMSHMOASIRVVAHLIEPFMMETS 480
Db 435 ALBAVWTLISRTNKYIDETAPWVLDKDEALRDOLASVMSHMOASIRVVAHLIEPFMMETS 494
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
Db 495 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 554
QY 541 EKWNPDVELKLNKDBIKFEDFKVEIRVAEVEKVSQVSGDKLLQFLDAGDGEDRQI 600
Db 555 EKWNPDVELKLNKDBIKFEDFKVEIRVAEVEKVSQVSGDKLLQFLDAGDGEDRQI 614
QY 601 LSGIAKYYPNEQELVGKKYQIVANLKPRKMKKKYVSQGMILSAEHDGKLTLLTVDPAPVN 660
Db 615 LSGIAKYYPNEQELVGKKYQIVANLKPRKMKKKYVSQGMILSAEHDGKLTLLTVDPAPVN 674
QY 661 GSVIG 665
Db 675 GSVIG 679

RESULT 3
H86723
methionine-tRNA ligase (EC 6.1.1.10) [imported] - Lactococcus lactis subsp. lactis (strai
N;Alternate names: methionyl-tRNA synthetase
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: H86723
R;Bolotin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrlic
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssi
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: H86723
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-662 <STO>
A;Cross-references: UNIPROT:Q9CHE0; GB:AE005176; PID:g12723712; PIDN:AAK04890.1; GSPDB:G
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: metS
C;Superfamily: methionyl-tRNA synthetase, dimer-forming
C;Keywords: ligase

Query Match 68.8%; Score 2392.5; DB 2; Length 662;
Best Local Similarity 68.6%; Pred. No. 2.3e-135;
Matches 459; Conservative 88; Mismatches 113; Indels 9; Gaps 7;
QY 1 MSEKNFYITPIYPSGKLIHGSAYTTIACDVLARYKRLMGVDVFLYGLDGHGKIQOK 59
Db 1 MTEKNFYITPIYPSGKLIHGSAYTTIACDVLARYKRLMGVDVFLYGLDGHGKIQOK 60
QY 60 KAEAGITPOAYVDGNAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAODDIYL 119
Db 61 KAEAGITPOAYVDGNAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAODDIYL 120
QY 120 GYSYSGHYSVDSDEFFTESQAEVFEDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDR 179
Db 121 LGKYAGHYSVDSDEFFTESQAEVFEDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDR 180
QY 180 RLVSEFFKAHPEFTTPDGRNLEMLNFIPEGLDLAVSRITFTTGVGPVPSNPKHVVVWIDA 239
Db 181 WLSQYIDEPDFIPEVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEV 240
QY 240 ALLNATYALGYAODEHGNFDFKFWNGTVFMVKGDKILRFHSIYWPILLMMLDKVLPDLIAH 299
Db 241 ALLNATYALGYAODEHGNFDFKFWNGTVFMVKGDKILRFHSIYWPILLMMLDKVLPDLIAH 298

```

QY 300 HGFWKDGKSKGNVYVPEMLVERYGIDPLRYLMLRNLPVSGDGTTPEDYVGRINY 359
DB 299 HGWLLKDGKSKGNVYVPEMLIBRYGLDAVRYLMLRAISFGQDGIPTPDEFVGRINF 358
QY 360 ELANDLGNLNRVSMINKYDFGQIPAYVEGV--TEEDHVLAEVAEKSIADFTHEAVDY 418
DB 359 DLANDLGNLNRVSMINKYNDGKIEA--TGVSTEDASLEEVETISHFHKAMDKPEF 416
QY 419 PRALAEVWTLISRTNKYIDETAPWVLDKDBALRDQLASVMSHWQASIRVVAHLIEPMMME 478
DB 417 NVRALADVWTLISRTNKYIDETAPWVLAKEBDDKAKLNNVLYHLAENLRITAGALLQPFMRA 476
QY 479 TSRAVLTQIGLEBS--SLENLSLADFPADVTVVAKGTPIFPRLNMEETIAYIKQOMEGNK 537
DB 477 TSKIEFQGNDRSFSLENLSFG--YGFTHPVVAKGQPIFPRLDVEEVAYIKLQWAGG- 534
QY 538 PAYEKEWNPDEVELKLNKDEIKFEEDPKVEIRVAEYKESKVEGSKLQFLRDAGDGB 597
DB 535 VLPEKEWNPDEVELNLTLPOIKEDDEKTELKVAEVEPEVBSDKLRFKLDAGDSBP 594
QY 598 RQILSGIAKYRNEOELVGVKQVIVANLKPRKMKKYSQGMILSAEHGDKLTLLTVDA 657
DB 595 RQILSGIAQYPNEQELVGVKQVIVANLKPRKMKKYSQGMILSAEFGDKLSVLTVD 654
QY 658 VPNGSVIG 665
DB 655 VPAGSLIG 662

RESULT 4
AB1097
methionyl-tRNA synthetase [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB1097
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok C.; Schluter, T.; Simeos, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A.; Title: Comparative Genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1097
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-664 <GLA>
A:Cross-references: UNIPROT:Q9YAP2; GB:NC_003210; PID:G16409534; GSPDB:
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: mets
C:Superfamily: methionyl-tRNA synthetase, dimer-forming

Query Match 60.7%; Score 2113; DB 2; Length 664;
Best Local Similarity 60.7%; Pred. No. 1.2e-118;
Matches 407; Conservative 101; Mismatches 144; Indels 18; Gaps 8;

QY 242 LNYATAGYAQDEHGNFDKFWNGTGFVFNVGKDILRFHSIYWPILLMMLDVKLDPRLIAHG 301
DB 244 SNYITAGYNTDNTDKFKYWPADV-QIVGKEIVRFHTIYWPIMLMALDPLPKKVPFGHG 302
QY 302 WFWKDGKSKGNVYVPEMLVERYGIDPLRYLMLRNLPVSGDGTTPEDYVGRINYEL 361
DB 303 WILKDGKSKGNVYVPEMLIDRYGLDALRYLLREVPEFGSDGLFTPEDFVDRVNF 362
QY 362 ANDLGNLNRVSMINKYDFGQIPAYVEGVTEFPHVLAEVAEKSIADFTHEAVDPRA 421
DB 363 ANDLGNLNRVSMINKYDFGQIPAYQGNVTEFPHVLAEVAEKSIADFTHEAVDPRA 422
QY 422 LEAVWTLISRTNKYIDETAPWVLDKDBALRDQLASVMSHWQASIRVVAHLIEPMMETS 481
DB 423 LNOQLSLISRTNKYIDETAPWALAKEDKRTSLASVTHLAENLRIITAVLLOFLTP 482
QY 482 AVLTOIGLEEVSSLENL-----SLADFPADVTVVAKGTPIFPRLNMEETIAYIKQOM 534
DB 483 EIFLQGLQE---ENLKKWDSIYGYGIEPAGTIVVAKGTPIFPRLDVAEVEVTIQDEMK 538
QY 535 GNKPAVEKWNPDVELKLNKDEIKFEEDPKVEIRVAEYKESKVEGSKLQFLRDAGD 594
DB 539 GSAPAPABEE--VAEVE-ALETPQIGIEDFKIDLRVAEVKQDVKKADKLLCFOLD 595
QY 595 GEDRQILSGIAKYRNEOELVGVKQVIVANLKPRKMKKYSQGMILSAEHGDKLTLLTV 654
DB 596 GKLRQVLSGIAEFYQPE-ELIGKVVIVVSNLKPVK-LRGLMSEGMILSGEKDKLSVIEA 653
QY 655 DPVPPNGSVI 664
DB 654 SSALPNGAKV 663

RESULT 5
AB1459
methionyl-tRNA synthetase [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB1459
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simeos, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A.; Title: Comparative Genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1459
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-664 <GLA>
A:Cross-references: UNIPROT:Q92P90; GB:AL592022; PID:G16412635; GSPDB:G
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: mets
C:Superfamily: methionyl-tRNA synthetase, dimer-forming

Query Match 60.6%; Score 2107; DB 2; Length 664;
Best Local Similarity 60.6%; Pred. No. 2.7e-118;
Matches 406; Conservative 101; Mismatches 145; Indels 18; Gaps 8;

QY 3 EKN-FYITPTIYPSGKLHIGSAYTTIACDVARYKRLMGYDVLYLTGLDEHGQKIQA 61
DB 4 EKNTFYITPTIYPSGKAHIGHAYTTVAGDAMARYKRLMGYDVLYLTGLDEHGQKIQA 63

QY 3 EKN-FYITPTIYPSGKLHIGSAYTTIACDVARYKRLMGYDVLYLTGLDEHGQKIQA 61
DB 4 EKNTFYITPTIYPSGKAHIGHAYTTVAGDAMARYKRLMGYDVLYLTGLDEHGQKIQA 63

QY 122 EYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSSEYFLRLSKYQDR 181
DB 124 EYEGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSSEYFLRLSKYQDR 183
QY 182 VEFKKAHPFITPDGRNLMNFIPLGDLAVSRFTTFWGVVPSNPKHVVVYVWIDAL 241
DB 184 VEYNGHPEPILPESRKNNMINFIKPLGDLAVSRFTTFWGVVPSNPKHVVVYVWIDAL 243

QY 62 BEAGITPOAYVDGMVGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 121
DB 64 KERGISEQYVDIAEGFQELWKLKLEISNTDFIRTTQDRHKTSVEKIFEQLLEGGDIYL 123
QY 122 EYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSSEYFLRLSKYQDR 181
DB 124 EYEGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSSEYFLRLSKYQDR 183

QY 182 VFFPKAHPFPIIPDGLNEMLNENFIEPGLDIAVSRRTFTWGVVPVSNPKHVVYWDAL 241
 DB 184 VYYNSHPFIFLPSKKNEMINNFIPGLDIAVSRRTFTWGVVPVSNPKHVVYWDAL 243
 QY 242 LNVATAGYAQDEHGNFDFKFMGTVPFHMVGKDIILFHSIYWPILMLDVLKPLDRLIAHG 301
 DB 244 SNYITAGYNTDNDKFKYKWPADV-QVIGKEIVAFHTIYWPILMALDVLKPLDRLIAHG 302
 QY 302 WFMKDGKMSKSGNVVPEMLVRYGLDPLRYLWRNLVPGSDGTTFTPPEDYVGRINVEL 361
 DB 303 WILMKDGKMSKSGNVVPEMLVRYGLDPLRYLWRNLVPGSDGTTFTPPEDYVGRINVEL 362
 QY 362 ANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSTADFTHTHEAVDYPRA 421
 DB 363 ANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSTADFTHTHEAVDYPRA 422
 QY 422 LSAVNTLSRTNKYIDETAPVWLDKDEALRDOLASVMHQAQIRVAHLIEPPNMTSR 481
 DB 423 LNLQSLISRTNKYIDETAPVWLDKDEALRDOLASVMHQAQIRVAHLIEPPNMTSR 482
 QY 482 AVLTQIGLEEVSSLENL-----SLADFPADVTVAKGTPFPEPLANNEEETIAYIKEOME 534
 DB 483 EIPQLQGLQEQ-----ENLKWDSIYGVGEIPFEGTTPVVKGTTPFPRDADAEVETIYQDEMK 538
 QY 535 GNKPAVEKEWNPDEVELKLNKDEIKFDPKVEIRVAEVEKSVKVEGSDKLLQFLDAGD 594
 DB 539 GSAPAPAE--TAEVE-ALETQIGIEDFDKIDLAEVAEKVQVQVKKADKLLCQLDLIGE 595
 QY 595 GEDRQLISGIAYKYPNEQELVKKVQVIVANLPRKMMKKYVSGMILSAEHGDKLTLTV 654
 DB 596 GKLRQVLSGIABFYQPE-ELIGKKVIVVSNLKPVK-LRGLMSGMLSGEXDKLSVIEA 653
 QY 655 DPAVPNGSVI 664
 DB 654 NSALPNGAKV 663

RESULT 6

S66067
 methionine-tRNA ligase (EC 6.1.1.10) mets - Bacillus subtilis
 N/Alternate names: methionyl-tRNA synthetase mets
 C/Species: Bacillus subtilis
 C/Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 R/Ogasawara, N.; Nakai, S.; Yoshikawa, H.
 C/Accession: S66067; E69657
 DNA Res. 1, 1-14, 1994
 A/Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis chrom
 A/Reference number: S65967; MUID:96051385; PMID:7584024
 A/Accession: S66067
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-664 <OGA>
 A/Cross-references: UNIPROT:P37465; EMBL:D26185; NID:G467326; PIDN:BAA05273.1; PID:G4674
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
 R/Kunst, P.; Ogasawara, N.; Wosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Bruchi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
 A.; Ehrlich, S.D.; Emerson, P.T.; Entlan, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koester, P.; Konings, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
 A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
 A/Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, K.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A/Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A.
 A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A/Reference number: A69580; MUID:98044033; PMID:9384377
 A/Accession: E69657
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA

A/Residues: 1-664 <KUN>
 A/Cross-references: GB:Z99104; GB:AL009126; NID:G2632267; PIDN:CAB11814.1; PID:G2632305
 A/Experimental source: strain 168
 C/Genetics:
 A/Gene: mets
 C/Superfamily: methionyl-tRNA synthetase, dimer-forming
 C/Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 56.4%; Score 1963.5; DB 2; Length 664;
 Best local similarity 57.8%; Pred. No. 1e-109;
 Matches 388; Conservative 106; Mismatches 162; Indels 15; Gaps 10;

QY 1 MSKN--FITTPIYPSGKLIHGSAYTTIACDLARYKRLMGDYVFLTGLDEHGKQIQ 58
 DB 1 MPQENNTFTTPIYPSGKLIHGSAYTTIACDLARYKRLMGDYVFLTGLDEHGKQIQ 60
 QY 59 QKAEAGITPQAVDGNVGVKELWQLLDISYDKFIRTTDDTHEKVAQVPELLAQDDI 118
 DB 61 QKAEQENITPQEVYDRAADIQKLMKQLKQLEISNDPFIITTEKHVKVIEKVFQKLDNGDI 120
 QY 119 YLGEYSGWYSVSDDEEFTTSQLAEVPDEAGNVGTGIAP-SGHEVEWSESEYFRLRSKY 177
 DB 121 YLDEYEGWISIPDETYYTITQLVDIRNEKGEVIGKSPDSGHPVELIKEESYFFFMGY 180
 QY 178 QDRLVEFFKAHPFPIIPDGLNEMLNENFIEPGLDIAVSRRTFTWGVVPVSNPKHVVYWD 237
 DB 181 ADSLLKYBENPTFIQPSKKNEMINNFIPGLDIAVSRRTFTWGVVPVSNPKHVVYWD 240
 QY 238 IDALLNVATAGYAQDEHGNFDFKFMGTVPFHMVGKDIILFHSIYWPILMLDVLKPLDRL 297
 DB 241 IDALFNLTALGVYDTEDELYKYPADV-HLVGREIVRFHTIYWPILMALDVLKPLDRL 299
 QY 298 IAHGWFVMDKMGKSGKSGNVVPEMLVRYGLDPLRYLWRNLVPGSDGTTFTPPEDYVGR 357
 DB 300 FAHGWLMLKDGKMSKSGNVVPEMLVRYGLDPLRYLWRNLVPGSDGTTFTPPEDYVGR 359
 QY 358 NYELANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSTADFTHTHEAVD 417
 DB 360 NYDLANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSTADFTHTHEAVD 419
 QY 418 YPRALAEVNTLSRTNKYIDETAPVWLDKDEALRDOLASVMHQAQIRVAHLIEPPNMTSR 477
 DB 420 FSVALTSLQVLSRTNKYIDETAPVWLDKDEALRDOLASVMHQAQIRVAHLIEPPNMTSR 479
 QY 478 ETSRAVLTIQGL--ESVSSLENLSADFPADVTVAKGTPFPEPLANNEEETIAYIKEOME 535
 DB 480 KTEPKMFEQLGITDESILKAWDSITAFQGLKD-TKVQKGEPLFPRLAEAEIEIAYIKMQG 538
 QY 536 NKPAVE--KEMNPDEVELKLNKDEIKFDPKVEIRVAEVEKSVKVEGSDKLLQFLDAG 593
 DB 539 SAPAKETKEEPEQEVN---RLPEITIDFMDVELVAEVEIAEPVKKADRLKLLQFLDAG 595
 QY 594 DGEDRQLISGIAYKYPNEQELVKKVQVIVANLPRKMMKKYVSGMILSAEHGDKLTLTV 653
 DB 596 -PEKQVWVSGIAKHVYTP-ELVGGKLVCTNLKPVK-LRGLMSGMLSGEXDKLSVIEA 652
 QY 654 VDPAPVNGSVI 664
 DB 653 IDQSLPKGTRI 663

RESULT 7

E83656
 methionyl-tRNA synthetase mets [imported] - Bacillus halodurans (strain C-125)
 C/Species: Bacillus halodurans
 C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
 A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A/Reference number: A83650; MUID:20512582; PMID:11058132
 A/Accession: E83656
 A/Status: preliminary

[illegible]

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: A97268
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-644 <MUC>
 A:Cross-references: UNIPROT:Q97EWS; GB:AE001437; PIDN:AAK80932.1; PID:g15026047; GSPDB:G000000000
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC2991
 C:Superfamily: methionyl-tRNA synthetase, dimer-forming

Query Match 47.2%; Score 1640.5; DB 2; Length 644;
 Best Local Similarity 48.3%; Pred. No. 28-90;
 Matches 322; Conservative 125; Mismatches 193; Indels 27; Gaps 10;

QY 1 MSEKNTTPTIYPSGKHLHGSAYTTIACDVLARYKLMGYDYFYLTGLDEHGGKIQK 59
 DB 1 MEKKTFTYTTIYPSGKHLHGSAYTTIACDVLARYKLMGYDYFYLTGLDEHGGKIQK 60
 QY 61 AEEAGITPOAYVDGMAVGKELWQLDLSYDKFIRTTDDYHEKVAQVAFERLLAQDDIYL 119
 DB 61 IAEKGIITPKAYVDELVAGIKLWKNMNTSYDKFIRTTDEEHVKAQVQKIVKFFDNGDIY 120
 QY 120 LGEYSWYSVDEEFTESQLAEVFDEAGNVTGGIAPSGHEVWVSSEYFLRLSKYQDR 179
 DB 121 KSAVEGWYCTPCESFYTERQL-----VDGKPCDCGRPEVKKEESYFFRMSKYVDR 170
 QY 180 LVEFFPKAHEPTTPGRLEMLRNFIKPLEDLAVSRFTFTWGPVPSNPKHVVVWIDA 239
 DB 171 LLQYEEENPDFIOPESRKNEMNINFIKPLEDLAVSRFTFTWGPVPSNPKHVVVWIDA 230
 QY 241 LLNAYATAGYAOQDEHGNFDFKFWNGVFMVGVKDIIRFHSIYWPILLMLMDLKLDRLLIAH 300
 DB 231 LANYITLALGYGTDNDEKRYKWPADV-HLVGKEIVRFHTIYWPIMMALGLPLPKVFGH 289
 QY 301 GWFVWKGMSKSGNVVYPEMLVERYGDLPLRYLYLARNLPVSGDGTTPEDYVGRINYE 360
 DB 290 GWLLMKDGMKSGNVVDPVMIIDRYGLDALRYLLREVPFGSDGVPTEGFTIERINYD 349
 QY 361 LANDLGNLLNRVTSMINKYFDGQIPAYVEGVTPEFPHVLAERVAEKSIAADFHTHMEADVPR 420
 DB 350 LANDLGNLLNRVTSMINKYFDGQIPAYVEGVTPEFPHVLAERVAEKSIAADFHTHMEADVPR 409
 QY 421 ALBAVNTLISRTNKYIDETAPWLDKDBALRDQLASVMSHWQASIRVVVAHLIEPFMMETS 480
 DB 410 ALSAVWQLIGRTNKYIDETAPWLDKDBALRDQLASVMSHWQASIRVVVAHLIEPFMMETS 469
 QY 481 RAVLTQIGLEEVSSLENLSLADFP--ADVTVAKGTPIFPRLNMBEETAYIKQMEGNKP 538
 DB 470 ERIFTQIGISDRSLKENDSLYDFGLIPGTNVQKGEPLFPRLDICEVEYIYKAEHQGGKP 529
 QY 539 AVEKEWNPDEVELKLNKDEIKPEDFKVEIRVAEVEKVSKEGSKLQFLDAGDGEDR 598
 DB 530 AEAAK----EEQAAAEBSIDDPKAVDLRVAEVVQPERMKNADKLLQLDLG-GEKR 584
 QY 599 QILSGIAKYPNBQELGVKKVQIVANLKPMMKKVVSQGMIL--SAEHGDKLTLLTVDP 658
 DB 585 QVLSGIAEYKPE-ELIGKKVIVANLKPMMKKVVSQGMIL--SAEHGDKLTLLTVDP 642
 QY 659 PNGSVI 664
 DB 643 PNGTKI 648

RESULT 10
 A97268
 methionyl-tRNA synthetase [imported] - Clostridium acetobutylicum
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
 C:Accession: A97268
 R:Noelling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 J. Bacteriol. 183, 4823-4838, 2001

A:Reference number: S16682; MUID:91305115; PMID:1852609
 A:Accession: S16682
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-649 <MEC>
 A:Cross-references: UNIPROT:P23920; EMBL:X57925; NID:g39988; PIDN:CAA40999.1; PID:g39988
 C:Genetics:
 A:Gene: mets
 C:Superfamily: methionyl-tRNA synthetase, dimer-forming
 C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

Query Match 55.2%; Score 1919; DB 2; Length 649;
 Best Local Similarity 55.9%; Pred. No. 4.5e-107;
 Matches 372; Conservative 103; Mismatches 171; Indels 20; Gaps 7;

QY 1 MSEKNTTPTIYPSGKHLHGSAYTTIACDVLARYKLMGYDYFYLTGLDEHGGKIQK 60
 DB 1 MEKKTFTYTTIYPSGKHLHGSAYTTIACDVLARYKLMGYDYFYLTGLDEHGGKIQK 60
 QY 61 AEEAGITPOAYVDGMAVGKELWQLDLSYDKFIRTTDDYHEKVAQVAFERLLAQDDIYL 120
 DB 61 AOEKGVITQQYVDDIVAGIQELWRKLDLSYDDFIRTTQERHKKIVKEIFARLVEQGIYL 120
 QY 121 GEYSWYSVDEEFTESQLAEVFDEAGNVTGGIAPSGHEVWVSSEYFLRLSKYQDR 180
 DB 121 GEYEGWYCTPCESFYTERQL-----VDGNCPCDCGRPEVKKEESYFFRMSKYVDR 170
 QY 181 LVEFFPKAHEPTTPGRLEMLRNFIKPLEDLAVSRFTFTWGPVPSNPKHVVVWIDA 240
 DB 171 LLQYEEENPDFIOPESRKNEMNINFIKPLEDLAVSRFTFTWGPVPSNPKHVVVWIDA 230
 QY 241 LLNAYATAGYAOQDEHGNFDFKFWNGVFMVGVKDIIRFHSIYWPILLMLMDLKLDRLLIAH 300
 DB 231 LANYITLALGYGTDNDEKRYKWPADV-HLVGKEIVRFHTIYWPIMMALGLPLPKVFGH 289
 QY 301 GWFVWKGMSKSGNVVYPEMLVERYGDLPLRYLYLARNLPVSGDGTTPEDYVGRINYE 360
 DB 290 GWLLMKDGMKSGNVVDPVMIIDRYGLDALRYLLREVPFGSDGVPTEGFTIERINYD 349
 QY 361 LANDLGNLLNRVTSMINKYFDGQIPAYVEGVTPEFPHVLAERVAEKSIAADFHTHMEADVPR 420
 DB 350 LANDLGNLLNRVTSMINKYFDGQIPAYVEGVTPEFPHVLAERVAEKSIAADFHTHMEADVPR 409
 QY 421 ALBAVNTLISRTNKYIDETAPWLDKDBALRDQLASVMSHWQASIRVVVAHLIEPFMMETS 480
 DB 410 ALSAVWQLIGRTNKYIDETAPWLDKDBALRDQLASVMSHWQASIRVVVAHLIEPFMMETS 469
 QY 481 RAVLTQIGLEEVSSLENLSLADFP--ADVTVAKGTPIFPRLNMBEETAYIKQMEGNKP 538
 DB 470 ERIFTQIGISDRSLKENDSLYDFGLIPGTNVQKGEPLFPRLDICEVEYIYKAEHQGGKP 529
 QY 539 AVEKEWNPDEVELKLNKDEIKPEDFKVEIRVAEVEKVSKEGSKLQFLDAGDGEDR 598
 DB 530 AEAAK----EEQAAAEBSIDDPKAVDLRVAEVVQPERMKNADKLLQLDLG-GEKR 584
 QY 599 QILSGIAKYPNBQELGVKKVQIVANLKPMMKKVVSQGMIL--SAEHGDKLTLLTVDP 658
 DB 585 QVLSGIAEYKPE-ELIGKKVIVANLKPMMKKVVSQGMIL--SAEHGDKLTLLTVDP 642
 QY 659 PNGSVI 664
 DB 643 PNGTKI 648

RESULT 11
 E72297
 methionine-tRNA ligase (EC 6.1.1.10) - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
 C:Accession: E72297
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,
 J. Bacteriol. 183, 4823-4838, 2001

C.M.

Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence
A:Reference number: A72200; YUID:99287316; PMID:10360571
A:Accession: E72297
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-629 <ARN>
A:Cross-references: UNIPROT:O33925; GB:AE001768; GB:AE000512; NID:g4981619; PIDN:AAD3616
A:Experimental source: strain MSB8
C:Genetics:
C:Superfamily: methionyl-tRNA synthetase, dimer-forming
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

```
Query Match      39.1%; Score 1359.5; DB 2; Length 629;
Best Local Similarity 42.4%; Pred. No. 1-2e-73;
Matches 284; Conservative 127; Mismatches 204; Indels 55; Gaps 14;

QY 6 FYITPIYPSGKLHIGSAYTTIACDVLYARKLMGYDVFLYLTGDEHGQK:OQKAEAG 65
DB 3 FYITPIYVNSPEPHIGSAYTTIADIIARYKRFMGYDVFLYLTGDEHGQK:VLAQAQAG 62
QY 66 ITPQAVVDMGVAVGKELWLLDISYDKFIRTTDDVHEKVVAVQVFERLLAQDDIYLGEYSG 125
DB 63 KDPQECDELAKEFKELWELKITNDYFIRTTDEMMKTVQFPAKCKENGVDYVIGIYK 122
QY 126 WYSVDEEFFTESQLAEVFRDEAGNVTGGIAPSGHE-----VEWVSEESYFLRLSK 176
DB 123 WYCVPCETFWNEDE-----VIKEGERFCPECKRPVKVVEEENYFFRLSK 167
QY 177 YQDLVEFFKAHPEFITPGRLEMLRNFIPEGLEDLAVSRFTTGWVPVPSNPHVTV 236
DB 168 YRDLKLKYEHPDFVEPFRNEMLK-ILEGLKDLSTRTTFKGVPMKDDPEHVIYV 226
QY 237 WIDALINYATALGYADDEHGNDFKFWNGTVFHMVGKDIIRFHSIYWPILLMLMDVKLPDR 296
DB 227 WVDALINYAIGYGNDE-MENKWPADL-HLIGKEINRFSIIWPAMLSVGLPLPK 284
QY 297 LIAHGFWMKDGMSKSGNVVYPEMLVERYGLDPLRYLYMLNLPVSGDGTTPDYGR 356
DB 285 VFAHGLTVNGQKISKSLGNAIDPRFFVKRYGNDVYVYLIIRDIMFGKDGDFSEERLVR 344
QY 357 INVELANDLGNLNRVTSMINKYFDGQI--PAYVEGVTEFDHVAEVAEKSIADPHTME 414
DB 345 LNSDLANDYGNLHRTAMIKKYFNGRLSPSAQEG---FDSWLKERFETKDAHEMD 401
QY 415 AVDYPALBAVTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLIEP 474
DB 402 SYRLTEALDKIWEFIADVKNKYFNDTKFWILGK-EGNMERLGTVLVNSLEAVEFKVALMTLP 460
QY 475 FWMETSRVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEBEETAYIKEQME 534
DB 461 VMPDTSSEVFRVRSFEKPSKSHLENGWVLPKGVSTVHGEPLFKKIDAKD-----FKVVE 516
QY 535 GNKPAVEKEWNPDEVELKLNKDEIKPEDPKVEIRVAEVEKYSKVEGSKLQFLRDLADG 594
DB 517 ----TVSAEQN-----AIIIDRFSKVDLAIAKVLAEKVPNSRKULRLIIDLGT 561
QY 595 GEDRQLLSGIKYPNEQBLVKKQVOIVANLKRPMKKYVSGMILSAEHGDKLTILTV 654
DB 562 -EKRQIVAGIAEHY-RPEELVGLKIVVANLKPAKLM-GIESQGMLLAASKGDTLLTV 618
QY 655 DPAPVPSGVI 664
DB 619 DGEITPGAKV 628
```

RESULT 12

F81356
methionine-tRNA ligase (EC 6.1.1.10) Cj0838c [imported] - Campylobacter jejuni (strain N
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jul-2004

C:Accession: F81356
R:Packhill, J.; Wren, B.W.; Mungall, K.; Kestley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: AB1250; YUID:20150912; PMID:10688204
A:Accession: F81356
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-628 <PAR>
A:Cross-references: UNIPROT:Q9P85; GB:AL139076; GB:AL111168; NID:g6968128; PIDN:CAB7310;
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
C:Superfamily: methionyl-tRNA synthetase, dimer-forming
C:Keywords: ligase

```
Query Match      34.2%; Score 1189.5; DB 2; Length 628;
Best Local Similarity 33.2%; Pred. No. 1.8e-63;
Matches 266; Conservative 132; Mismatches 209; Indels 71; Gaps 18;

QY 7 YITPIYPSGKLHIGSAYTTIACDVLYARKLMGYDVFLYLTGDEHGQK:OQKAEAGI 66
DB 3 YITPIYVNDVPHLGHAYTTIADTLARFVRLQGHETRFLTGDEHGQK:LEAAKLNS 62
QY 67 TPQAVVDMGVAVGKELWLLDISYDKFIRTTDDVHEKVVAVQVFERLLAQDDIYLGEYSGW 126
DB 63 TPQYADKISPEPKLWDEFEITYDIARTDTRHIEFIKAMFLKMWQKGDYKDEYEGH 122
QY 127 YSVSDEEFFTESQLAEVFRDEAGNVTGGIAPS-GHEVWVSEESYFLRLSKYQDLVEFF 185
DB 123 YCISCEFFTQSQJ-----INDCSPDCGKQTRILKESSYFFKLKYQKILQWY 172
QY 186 KAHEFITPGRLEMLRNFIPEGLEDLAVSRFTTGWVPV-----SNPKHVYVYVWDALL 242
DB 173 E-EKDPILPKNKNELI-NFVQGLKDLSTRTSFDWGIKLPQEIINDKHIIYVWLDAIF 230
QY 243 NYATALGYADDEHGNDFKFWNGTVFHMVGKDIIRFHSIYWPILLMLMDVKLPDLIAHG 302
DB 231 IYVSSLDLF--QNKGENAKFWPAHV-HLVGKDIIRFHAHYWPAFLMSVDLPLPKFIGA 287
QY 303 FVMKDGMSKSGNVVYPEMLVERYGLDPLRYLYMLNLPVSGDGTTPDYVGRINYELA 362
DB 288 WTKEGEMSKSGNVVYKPEVDVAYGSEAFRYFLREVPGNDGDFSENMLNINRAELS 347
QY 363 NDLGNLNRVTSMINKYFDGQI PAYVEGVTEFDHVAEVAEKSIADPHTMEAVDYPRAL 422
DB 348 NEFGNLLNRIIGMSTKYSQGNISK--EGVLKFYNAELNQAKEHLNLAVEFLENLQCNRYL 405
QY 423 EAVVTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRV-----VAHLIE 473
DB 406 EELFKALSVANLAISKYEPHSLIKEN-----KHEQANALVALCANILAKTSLLS 455
QY 474 PFMETSRVLTQIGLEEVSS-----LENLSLADFPADVTVVAKGTPIFPRLNMEEBEIA 527
DB 456 PTLPKSKQKVALANFETISSANVTKMILDN-ELLDFKANPC-----EALPKV-----EKA 505
QY 528 YIKQMEGNKPAVEKEWNPDEVELKLNKDEIKPEDPKVEIRVAEVEKYSKVEGSKLQ 587
DB 506 LLKQEI---KEEPKKEESP-----KIKIDDFAKIEIKVAKVLDCCQIESEKLLK 552
QY 588 FRLDAGDGEDRQLLSGIKYPNEQBLVKKQVOIVANLKRPMKKYVSGMILSAEHGDK 647
DB 553 FQELDDKEIRQVLSGIKYY-KASDLIGKQVCVISNLKKAIF-GHESDGMILSAKSGD 610
QY 648 KLTLLTVDPAPVPSGVI 665
DB 611 KLVLIAPPEQLVQNGSLVG 628
```

RESULT 13

C82922
methionyl-tRNA synthetase UUI97 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Feb-2003
C:Accession: C82922
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mir
A:Reference number: A62870
A:Accession: C82922
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-521 <GLA>
A:Cross-references: GB:AB002119; GB:AF222894; NID:G6899153; PIDN:AAF30604.1; GSPDB:GN001
A:Experimental source: serovar 3, biovar 1
C:Genetics:
A:Gene: mets; UU197
A:Genetic code: SGC3
C:Superfamily: methionyl-tRNA synthetase

Query Match	33.2%;	Score 1153.5;	DB 2;	Length 521;
Best Local Similarity	43.7%;	Pred. No. 1.9e-61;		
Matches 231;	Conservative 104;	Mismatches 181;	Indels 13;	Gaps 9;

QY 1 MSEKNFYTTIYPSGKHLGHSAYTTIACDVLARYKRLMGVDYFYLTLGLDGHQKIQOK 60
DB 2 LKQKFFISTPIYSSGNPHIGHAYTTIADVLARYKRLFGYDFVFTLGMDEHGKIQOK 61
QY 61 AEEAGITTPQAVDGMAGVVKELWQLLDISYDKFIRTTDDYHEKVVAQVFERLLAQDDIYL 120
DB 62 AFEENISPKALVDNRNIIPLNKLNLISFSKFIQTQMDHEESVQKVSFLYKQGIYL 121
QY 121 GYSGWYSVSEDEPTESQLAEVFRDEAGNVTGGIAPSGCHEVWVSSEYFLRLSKYQDR 180
DB 122 GOWTGYCVSEENTNP---AEIKSQ-DNIM--LCRMGHKLETKSEESYFYKMSDAQPF 175
QY 181 LVEPFKAPERTTPDGRNEMLRNFIPEGLEDLAVSRRTFTWGVPSNPKNHVVYVWIDA 240
DB 176 LKTYQNHFNFIIPERANENWNNFN-NLEDLSISRTFTDNGPIAENPKHVIYVWLDA 234
QY 241 LNYATALGYAODERGNDFKW--NGT-VFMVGKDILRFHSIYWPILLMLMDVLKPDRL 297
DB 235 LMNYLTATGYLSNNEELFKYWCNDNETEIVHLLSKEIARFHCIIYWPILNDLQIRFPSTI 294
QY 298 IAHGFVMDGKMSKGNVYVPEMLVERYGDLPLRYLMLNLPVSGDGTTPEDYVGRI 357
DB 295 LSHGWIITKEGKVSGLNVDNPVLDITDYGVDALRYIYIMADLSLFRDAIFSDNLITY 354
QY 358 NYELANDGLNLRVTSVMINKYFDGQIPAYVEGVTFDFHVLAEVAKSIADPFTHMEAVD 417
DB 355 NTQLANSYGNMISRTGLMKYRNIVPKYGVCLKNDEKLENLNKIELVQENINKYS 414
QY 418 YPRALEAVWTLISRTNKYIDETAPWLDKDEALRDLQASVMSHWQASIRVVAHLIEPFMM 477
DB 415 IDKALNCIQEILVEANKYIEDNKPKWLAKNQO-EQELDSLIVHLVKVIQVTTLLSPILI 473
QY 478 ETSRAVLTQIGSEYSSLENLSLADPPA-DVTVAKGTPIPRLNMEEE 525
DB 474 EGSKAVEQLNFE-SFLTASLASYDIFNHYKNDSEKPIARIIVEKQ 521

RESULT 14
H71867
methionine-tRNA ligase (EC 6.1.1.10) - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: H71867
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, B.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Malls, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: H71867
A>Status: preliminary

A:Molecule type: DNA
A:Residues: 1-656 <ARN>
A:Cross-references: UNIPROT:Q9ZKG9; GB:AE001525; GB:AE001439; NID:G4155533; PIDN:AA0654;
A:Experimental source: strain J99
C:Genetics:
A:Gene: metG
C:Superfamily: methionyl-tRNA synthetase, dimer-forming
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match	33.1%;	Score 1152.5;	DB 2;	Length 656;
Best Local Similarity	39.9%;	Pred. No. 3.1e-61;		
Matches 275;	Conservative 116;	Mismatches 227;	Indels 71;	Gaps 19;

QY 8 ITTPIYPSGKHLGHSAYTTIACDVLARYKRLMGVDYFYLTLGLDGHQKIQOKAEEAGIT 67
DB 6 ITTPIYVNDIPHIGHAVTTIADTLKKYTIQGSSEVFTGTDEHGQKIEQSARLNQS 65
QY 68 POAYVDGMAGVVKELWQLLDISYDKFIRTTDDYHEKVVAQVFERLLAQDDIYLGYSQWY 127
DB 66 PRAYADSIASIFKNQWDFNLDYDGFIRTTDSEHOKCVQNAFEIMFEGDIYKGYISY 125
QY 128 SVSDEEFFTESQLAEVFRDEAGNVTGGIAPSG-HEVEWVSSEYFLRLSKYQDRIVERFK 186
DB 126 CVSSEYCAVSKV---DNTDSKV--LCPDCLRETTLESESYFFKL SAYEKPILLEFYA 178
QY 187 AHPEFITPDGRNEMLRNFIPEGLEDLAVSRRTFTWGVVP---SNPKHVVYVWIDALLN 243
DB 179 KNEPAIPLPYRKNE-VTSFIEQGLDLSITRFSFEGWIPLPKKNMDPKHVVVWIDALLN 237
QY 244 YATALGY---AODEHGNFDKFWNGTVFMVGKDILRFHSIYWPILLMLMDVLKPDRLJAH 300
DB 238 YASALGYLNLGNLKNWAHER-----ARHVGKDILRFHAIYWPFAFLMSLNLPLPKQLCVH 292
QY 301 GFMVKMGKMSKGNVYVPEMLVERYGDLPLRYLMLNLPVSGDGTTPEDYVGRINYE 360
DB 293 GWWTIEGVMSKSLGNVLDAQKLAWEYGLIEURLYFLREVFGQDGFSSKALVERINAN 352
QY 361 LANDGLNLRVTSVMINKYFD---GOIPA VEGVTFDFHVLAEVAKSIADPFTHMEA 415
DB 353 LNDLGNLNRLLGHAKKIFNYSLSKTKITAYYPKELEKAKHQLDNANSFVPKQLH--- 409
QY 416 VDYPRALEAVWTLISRTNKYIDETAPWLDKQ-----EALRDLQASVMSHWQASIRVVA 469
DB 410 ----KALEELFNIDYFLNKLIAKEEPFWLHKHKNSEKLEALLSLIANTL--LQSSFLLYA 463
QY 470 HLIETPMETS--RAVLTQIGLE--EVSSENLSIADPPADVTVVAKGTPIPRLNMEEE 524
DB 464 FMPKSAMKLASAPRVEITPNNYERFFKAKLQDMVLQD-----TEPLFSKIEKIE 513
QY 525 ETAYIK--EQMEGNKPAV-----EKWNPDEVELKLNKDEIKFEDFKVEIRVAEVRK 575
DB 514 KIEKIEKIEKIEGEAEALAEKAEKKEKAPTOE-----NVISIEDFKVEIKVGLIKE 568
QY 576 VSKVSGDKLQFRLDAGDGEORILSGIATKYPNEQELVGKQVQIVANLKRKMKKVV 635
DB 569 AORIEKSNKLLRLKVDLGENRLQIISGIALDYEPB-SLVGQMVVAVNLKPAKLMGE-M 626
QY 636 SQGMILSAEHGDKLTLTVDPAVPGSVI 664
DB 627 SEGMLAVRDNDNLALISPTREKIAGSLI 655

RESULT 15
A64572
methionine-tRNA ligase (EC 6.1.1.10) - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C:Accession: A64572
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khairak, H.G.; Glodek, A.; McKenney
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: A64572

A>Status: Preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-650 <TOM>

A:Cross-references: UNIPROT:P56127; GB:AB000557; GB:AB000511; NID:G2313514; PIDN:AAD0748

C:Superfamily: methionyl-tRNA synthetase, dimer-forming

C:Keywords: ligase

```
Query Match      33.0%; Score 1148.5; DB 2; Length 650;
Best Local Similarity 39.6%; Pred. No. 5.3e-61;
Matches 268; Conservative 120; Mismatches 236; Indels 53; Gaps 16;

Qy      8 ITTPIYVPSGKLIHGSAVTIACDVLARYKELMGVDVYLTGLDGHGKIQQKAEERAGIT 67
Db      6 ITTPIYVNDIPHCHAYTHIADTLKYYTLQGEVFFLTGDEHGCKIEQEARLNQS 65

Qy      68 PQAYVDMAGVVKELMQLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYLGESGWY 127
Db      66 PKAYADGISITFKQDWFNLDYDGFIRTTDSEHOKCVQNAFEIMPEKGDYKGAISGY 125

Qy      128 SVSDEEFFTSQLAEVRDEAGNVGTGIIAPSG-HEVEWVSEESYFLRLSKYQDRLYEFPK 186
Db      126 CVSCESYCALSK-ADNTNDKV-----LCPDCURETTLLSEESYFFRLSAYEKPLDLYA 178

Qy      187 AHPEFITPDGRLNEMLRNFIEPGLDLAVSRITFTMGVPVP---SNPKHVYVYVMDALLN 243
Db      179 KNPEAILPVYRKNE-VTSFIEQGLDLSITRTSPFEMGIPLPKQWDPKHVVYVWLDAALLN 237

Qy      244 YATALGYAQDEHGNFDFWNGTVFHMVGKDIILRHFSIYWFILLMLDVKLPDRLIAHGWP 303
Db      238 YASALGYNLNDLNKQAHF--ECARHIVGKDIRFHAIYPAFLMSLNLPLFKQLCVHGW 295

Qy      304 VMKDGKMSKSGNVVYVPEMLVERGYLDPLRYLMLNLPVSGDGTFTPTEDYVGSINVELAN 363
Db      296 TIEGVNMSKSLGNVLDQAQKAMEYGEELRYFLRLREVFGQDGDGFKKALIERINANLNN 355

Qy      364 DLGNLNRVTSMINKYFD----GOIPAYVEGVTEFDHVLAEVAEKSIADFTHTMEAVDY 418
Db      356 DLGNLNRLLIGMAKCYFNHSLKSTKITAYYSKELEKVKHQILDNANSFVPKQQLH----- 409

Qy      419 PRALAEVTLISRTNKVIDETAPWLDKDEALRDQDLASVMHMQASIRVVAHLIEPFM-- 476
Db      410 -KALEELFNYYDFLNKLIKAEKPPWLFKKNNE-SEKLEALLSLIANALLQSSFLLYAFMPK 467

Qy      477 --METSRAVLTOIGLE-----EVSSLENLSLADFPADVTVAKGTPIPERLANNEEYAY 528
Db      468 SAVKLANAFNTEITPDNYERFFKAKKLQDMILQD-----TEPLFSKM---EKIEK 514

Qy      529 IKEQMEGNKPAVEKEMNPDEVELKLNKDE-IKPEDFKVEIRVAEYKSVKVEGSKLLQ 587
Db      515 TERAGEASPEKNEKEKDAKAPLQENYIGIEDPKKVEIKVGLIKEAQRIEKNKLLR 574

Qy      588 FRLDAGGEDROILSGIAKYYPNEQELVGGKQIVANLKPRQMKYVSGMILSAEHDG 647
Db      575 LKVDLGEGRLRQIISGIALDYEPF-SLVGMVCVAVNLKPAKLMBE-MSEGMITLAVRDS 632

Qy      648 KLTLLITVDPAVPNGSVI 664
Db      633 NLALISPTREKIAGSLI 649
```

Search completed: December 17, 2004, 22:00:18

Job time : 44 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2004, 21:42:30 ; Search time 200 Seconds
(without alignments)
1913.121 Million cell updates/sec

Title: US-10-606-607-2
Perfect score: 3479
Sequence: 1 MSEKNFYITPIYVPSGKLH.....DGKLTLLTVDPAPVNGSVIG 665

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3446	99.1	665	1 SYM_STRPN	Q9716 streptococc
2	2883	82.9	669	2 Q8DSW8	Q8DSW8 streptococc
3	2880.5	82.8	665	1 SYM_STRP3	P59080 streptococc
4	2879.5	82.8	665	1 SYM_STRP5	Q9A178 streptococc
5	2868.5	82.5	665	1 SYM_STRP8	Q8P298 streptococc
6	2732.5	78.5	665	2 Q8D7D0	Q8D7D0 streptococc
7	2732.5	78.5	665	2 Q8E3Z3	Q8E3Z3 streptococc
8	2322.5	68.8	682	1 SYM_LACLA	Q9CHE0 lactococcu
9	2288	65.8	669	1 SYM_ENTFA	Q837B3 streptococc
10	2230.5	64.1	673	1 SYM_OCEIH	P59079 oceanobacill
11	2201	63.3	683	1 SYM_LACPL	Q88297 lactobacill
12	2114	60.8	665	2 Q72FN6	Q724N6 listeria mo
13	2114	60.8	665	2 AAT02975	AAT02975 listeria
14	2113	60.7	664	1 SYM_LISMO	Q8YAF2 listeria mo
15	2107	60.6	664	1 SYM_LISIN	Q92190 listeria in
16	1990	57.2	658	2 Q74LI3	Q74LI3 lactobacill
17	1990	57.2	658	2 AAS08184	AAS08184 lactobacil
18	1963.5	56.4	664	1 SYM_BACSU	P37465 bacillus su
19	1953.5	56.2	660	1 SYM_BACHD	Q9KGK8 bacillus ha
20	1951.5	56.1	660	2 Q73FH0	Q73FH0 bacillus ce
21	1951.5	56.1	660	2 AAS38972	AAS38972 bacillus
22	1939.5	55.7	660	2 Q6HFX8	Q6HFX8 bacillus th
23	1936.5	55.7	660	1 SYM_BACCR	Q81JA8 bacillus ch
24	1936	55.6	660	1 SYM_BACJ1	Q6GJ11 staphylococc
25	1935.5	55.6	660	1 SYM_BACAN	.Q81W03 bacillus an
26	1935.5	55.6	660	2 AAT79115	AAT79115 bacillus
27	1931	55.5	657	1 SYM_STRAM	Q99WB3 staphylococ
28	1929	55.4	657	1 SYM_STRAM	Q8NY00 staphylococ
29	1929	55.4	657	2 Q6GBZ8	Q6GBZ8 staphylococ
30	1922.5	55.3	656	1 SYM_STRAP	Q8CQJ3 staphylococ
31	1919	55.2	643	1 SYM_BACST	P23920 bacillus st

32	1701	48.9	638	1 SYM_THETN	Q8RDD1 thermoanaer
33	1664	47.8	641	1 SYM_CLOTE	Q899D9 clostridium
34	1651.5	47.5	645	1 SYM_CLOPE	Q8XNG1 clostridium
35	1640.5	47.2	644	1 SYM_CLOAB	Q97EW5 clostridium
36	1499	43.1	636	2 Q7P484	Q7P484 fusobacteri
37	1484	42.7	636	1 SYM_FUSNN	Q8RES7 fusobacteri
38	1421.5	40.9	418	2 Q7WSP6	Q7WSP6 bacillus ce
39	1417.5	40.7	416	2 Q7WSQ0	Q7WSQ0 bacillus ce
40	1407.5	40.5	416	2 Q7WSP5	Q7WSP5 bacillus ce
41	1387.5	39.9	411	2 Q7WSP7	Q7WSP7 bacillus ce
42	1382.5	39.7	410	2 Q7WSP8	Q7WSP8 bacillus ce
43	1359.5	39.1	629	1 SYM_THEME	Q33925 thermotoga
44	1340.5	38.5	400	2 Q7WSQ1	Q7WSQ1 bacillus ce
45	1329.5	38.2	397	2 Q7WSP9	Q7WSP9 bacillus ce

ALIGNMENTS

RESULT 1
ID SYM_STRPN STANDARD; PRT; 665 AA.
AC Q97RL6,
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Methylonyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase) (MetRS).
DE Name=metG; Synonyms=metS; OrderedLocustNames=SP0788, spr0696;
OS Streptococcus pneumoniae, and
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313, 171101;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D., Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J., Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D., Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D., Holtzapple E.K., Khouri H.M., Wolf A.M., Uterback T.R., Hansen C.L., McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T., Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C., Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus pneumoniae."
RL Science 293:498-506(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-255 / R6;
RX MEDLINE=21429245; PubMed=11544234;
DOI=10.1128/JB.193.19.5709-5717.2001;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S., DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geisinger C., Gilmore R., Glass J.S., Kuo H., Kraft A.R., Lagace R.E., LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P., McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Niclas T.I., Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P., Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G., Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L., Glass J.I.;
RA "Genome of the bacterium Streptococcus pneumoniae strain R6.";
J. Bacteriol. 183:5709-5717(2001).
RL J. Bacteriol. 183:5709-5717(2001).
CC -!- FUNCTION: Is required not only for elongation of protein synthesis but also for the initiation of all mRNA translation through initiator tRNA(iMet) aminoacylation (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP + diphosphate + L-methionyl-tRNA(Met).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.

CC MetG subfamily 2B.
 CC -!- SIMILARITY: Contains 1 tRNA-binding domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE007385; RAK74925.1; -;
 CC EMBL; AE008446; AAK99500.1; ALT_INIT.
 CC FIR; D95091; D95091.
 CC HSSP; Q9V011; 1MKH.
 CC TIGR; SP0788; -;
 CC HAMAP; MF 01228; fused; 1.
 CC InterPro; IPR004495; MetG Cterm.
 CC InterPro; IPR008224; MetRS dimerising.
 CC InterPro; IPR002300; tRNA-synt_1a.
 CC InterPro; IPR001412; tRNA-synt_1.
 CC InterPro; IPR002304; tRNA-synt_met.
 CC InterPro; IPR009080; tRNA-synt_1a_bind.
 CC InterPro; IPR002547; tRNA_bind.
 CC Pfam; PF01333; tRNA-synt_1; 1.
 CC Pfam; PF01588; tRNA_bind_1.
 CC PIRSF; PIRSF001528; MetRS dimerising; 1.
 CC PRINTS; PR01041; TRNASYNTHMET.
 CC TIGRFAMS; TIGR00398; metG; 1.
 CC TIGRFAMS; TIGR00399; metG_C term; 1.
 CC PROSITE; PS00178; AA TRNA LIGASE I; FALSE_NEG.
 CC PROSITE; PS00896; TRBD; 1.
 CC Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
 KW Protein biosynthesis; RNA-binding; tRNA-binding.
 KW SITE 13 23 "HIGH" region.
 FT SITE 309 313 "KMSKS" region.
 FT DOMAIN 562 665 tRNA-binding.
 FT BINDING 312 312 ATP (By similarity).
 SQ SEQUENCE 665 AA; 75651 MW; 7FAB60A5CC49A56D CRC64;
 Query Match 99.1%; Score 3446; DB 1; Length 665;
 Best Local Similarity 99.2%; Pred. No. 3.5e-198;
 Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYGLDGHGQKIQQK 60
 DB 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYGLDGHGQKIQQK 60
 QY 61 ABEAGITPQAYDGMAGVKELWQLDIDSYDKFIPTDDYHKKVAQVFPELLAODDIYL 120
 DB 61 ABEAGITPQAYDGMAGVKELWQLDIDSYDKFIPTDDYHKKVAQVFPELLAODDIYL 120
 QY 121 GEVSGNYSDEEFFTESQAEVFRDEAGNVGTGGIAPSGHEVVEVSESSYFLRLSKYQDR 180
 DB 121 GEVSGNYSDEEFFTESQAEVFRDEAGNVGTGGIAPSGHEVVEVSESSYFLRLSKYQDR 180
 QY 181 LVEFFKAHPEFITPDGRNLNMLNFIEPGLEDLAVSRITFTMGVVPVSPNKHVVVWIDA 240
 DB 181 LVEFFKAHPEFITPDGRNLNMLNFIEPGLEDLAVSRITFTMGVVPVSPNKHVVVWIDA 240
 QY 241 LNNYATGACADENGFNFKNVGTVMVGVKDIIRFHSIYWPILLMLDKVLPDLRLIAH 300
 DB 241 LNNYATGACADENGFNFKNVGTVMVGVKDIIRFHSIYWPILLMLDKVLPDLRLIAH 300
 QY 301 GFWFMKDGKMSKGNVYYPEMLVERYLGLDPLRYLYLMRNLPGVSGDTFTPEDYVGRINYE 360
 DB 301 GFWFMKDGKMSKGNVYYPEMLVERYLGLDPLRYLYLMRNLPGVSGDTFTPEDYVGRINYE 360
 QY 361 LANDLGNLNRVTSMNKYFDGQIPAYVEGVTFEFDHVLAEVAEKSIADPHTHMEADVPR 420
 DB 361 LANDLGNLNRVTSMNKYFDGQIPAYVEGVTFEFDHVLAEVAEKSIADPHTHMEADVPR 420
 QY 421 ALEAVVTLISRTNKYIDETAPVWLKDEALRDLQSLASVMSHMQASIRVVAHLIEPFMMETS 480

DB 421 ALEAVVTLISRTNKYIDETAPVWLKDEALRDLQSLASVMSHMQASIRVVAHLIEPFMMETS 480
 QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTTVAKGTPIFPELANMBEEIAY:KEQMEGNKPAV 540
 DB 481 RAVLTQIGLEEVSSLENLSLADFPADVTTVAKGTPIFPELANMBEEIAY:KEQMEGNKPAV 540
 QY 541 EKWNVDEVELKLNKDEIKFEDFDKVIIRVAEVEKVEKSGSKLLOPRLDAGDGEDROI 600
 DB 541 EKWNVDEVELKLNKDEIKFEDFDKVIIRVAEVEKVEKSGSKLLOPRLDAGDGEDROI 600
 QY 601 LSGIAKYYPNEQELVGKQVQIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVEN 660
 DB 601 LSGIAKYYPNEQELVGKQVQIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVEN 660
 QY 661 GSVIG 665
 DB 661 GSVIG 665
 RESULT 2
 Q8DSW8 ID Q8DSW8 PRELIMINARY; PRT; 669 AA.
 AC Q8DSW8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative methionyl-tRNA synthetase (EC 6.1.1.10).
 GN Names: Streptococcus mutans.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
 RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
 RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
 RA Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 RT pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 DR EMBL; AE014994; AAN59279.1; -;
 DR HSSP; Q9V011; 1MKH.
 DR GO; GO:0005737; Cytoplasm; IEA.
 DR GO; GO:0005524; F-ATP binding; IEA.
 DR GO; GO:0016874; Filigase activity; IEA.
 DR GO; GO:0004825; F-methionine-tRNA ligase activity; IEA.
 DR GO; GO:0000049; F-tRNA binding; IEA.
 DR GO; GO:0006431; P-methionyl-tRNA aminoacylation; IEA.
 DR GO; GO:0006412; P-protein biosynthesis; IEA.
 DR InterPro; IPR004495; MetG Cterm.
 DR InterPro; IPR002300; tRNA-synt_1a.
 DR InterPro; IPR002304; tRNA-synt_met.
 DR InterPro; IPR009080; tRNA-synt_1a_bind.
 DR InterPro; IPR002547; tRNA_bind.
 DR Pfam; PF01333; tRNA-synt_1; 1.
 DR Pfam; PF01588; tRNA_bind_1.
 DR PRINTS; PR01041; TRNASYNTHMET.
 DR TIGRFAMS; TIGR00398; metG; 1.
 DR TIGRFAMS; TIGR00399; metG_C term; 1.
 DR PROSITE; PS00896; TRBD; 1.
 KW Aminoacyl-tRNA synthetase; Complete proteome; Ligase.
 SQ SEQUENCE 669 AA; 75635 MW; 10B0B3ACE1DDCEE CRC64;
 Query Match 82.9%; Score 2883; DB 2; Length 669;
 Best Local Similarity 82.4%; Pred. No. 2.1e-164;
 Matches 553; Conservative 46; Mismatches 64; Indels 8; Gaps 5;
 QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYGLDGHGQKIQQ 59

Db 1 MTEKQPFYITPIYPSGKLHIGSAVTTIACDVLARYKEMMNDVFFYLTGLDHEGQKIQ 60
QY 60 KAEAGITPQAYVDGMAVGKELWQLDLSYDKFIRTTDDYHEKVAQVFERLLAQDDIY 119
Db 61 KSEAGITPQAYVDGMAVGKELWQLDLSYDKFIRTTDDYHEKVAQVFERLLAQDDIY 120
QY 120 LGYSGWYSVDSDEFFTESQLEAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYOD 179
Db 121 LGYSGWYSVDSDEFFTESQLEAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYAD 180
QY 180 RLVEFFKAMPERTTPGRNEMLRNIEPGLDVLAVSRITFTMGVVPVPSNPKHVYVWID 239
Db 181 RLVDFFFAHPDFIOPGRNEMLRNIEPGLDVLAVSRITFTMGVVPVPSNPKHVYVWID 240
QY 240 ALLNYATAGYAGDERHGNDFKPNWGVTFVFMVGNKDLIRFHSIYWPIILLMLDLVKLPDLRLFA 299
Db 241 ALLNYATAGYAGDERHGNDFKPNWGVTFVFMVGNKDLIRFHSIYWPIILLMLDLVKLPDLRLFA 300
QY 300 HGFVFMKDGKSKSGNVVYVPEMLVRYGLDPLRYLYMRLNPLVSGSDGTTTPEYVGRIN 359
Db 301 HGFVFMKDGKSKSGNVVYVPEMLVRYGLDPLRYLYMRLNPLVSGSDGTTTPEYVGRIN 360
QY 360 ELANDLGNLNRVTSMINKYFDQIPAYVEGVTFDFHVLAEVAESIAIDFHTMEAVDYP 419
Db 361 ELANDLGNLNRVTSMINKYFDQIPAYVEGVTFDFHVLAEVAESIAIDFHTMEAVDYP 420
QY 420 RALEAVWTLISRTNKYIDETAPWLDKDEALRQLASVMSHQAQSTVRVAHIEPFWMET 479
Db 421 RALEAVWTLISRTNKYIDETAPWLDKDEALRQLASVMSHQAQSTVRVAHIEPFWMET 480
QY 480 SRAVLTCIGLEEVSSLENISLADFPADVTVAAGTPIFFRLNMEEBIAIKQMEGNKPA 539
Db 481 SRAVLTCIGLEEVSSLENISLADFPADVTVAAGTPIFFRLNMEEBIAIKQMEGNKPA 539
QY 540 V-----EKEMNPDEVELKANKDEIKFEDFKVEIRVAEVEKVSQKLLQFRLDAGDG 595
Db 540 V-----EKEMNPDEVELKANKDEIKFEDFKVEIRVAEVEKVSQKLLQFRLDAGDG 595
QY 596 EDQILSGIAKYPNQEQLGKVKQVIVANLKPKKMKKVSQGMILSAEH-DGKLTLLTV 654
Db 596 EDQILSGIAKYPNQEQLGKVKQVIVANLKPKKMKKVSQGMILSAEH-DGKLTLLTV 654
QY 655 DPVAVNGSVIG 665
Db 655 DPVAVNGSVIG 665
RESULT 3
SYM_STRP3 STANDARD; PRT; 665 AA.
AC P59080;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)
DE (MetRS).
GN Names=metG; Synonyms=metS; OrderedLocusNames=SpyM3_0300, Spal1557;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133809; PubMed=12122206; DOI=10.1073/pnas.152298499;
RA Bares S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.P., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence."
RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).

[2]
RN SEQUENCE FROM N.A.
RP STRAIN=SSI-1 / Serotype M3;
RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution."
RL Genome Res. 13:1042-1055 (2003).
CC -!- FUNCTION: Is required not only for elongation of protein synthesis
CC but also for the initiation of all mRNA translation through
CC initiator tRNA (Met) aminoacylation (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA (Met) = AMP +
CC diphosphate + L-methionyl-tRNA (Met).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC MetG subfamily 2B.
CC -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE014142; AAW78907.1; -
CC EMBL; AP005145; BAC64652.1; ALT_INIT.
CC HSSP; P23395; 1A8H.
CC HAVAP; MF_01428; fused; 1.
CC InterPro; IPR004495; MetG_Cterm.
CC InterPro; IPR009224; MetRS_dimerising.
CC InterPro; IPR002300; tRNA-synt 1a.
CC InterPro; IPR001412; tRNA-synt 1.
CC InterPro; IPR002304; tRNA-synt_met.
CC InterPro; IPR009080; tRNA-synt_1a_bind.
CC InterPro; IPR002547; tRNA_bind.
CC Pfam; PF00133; tRNA-synt 1; 1.
CC Pfam; PF01588; tRNA_bind; 1.
CC PIRSF; PIRSF001528; MetRS_dimerising; 1.
CC PRINTS; PR01041; TRNASYNTHMET.
CC TIGRFAMs; TIGR00398; metG; 1.
CC TIGRFAMs; TIGR00399; metG_Cterm; 1.
CC PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
CC PROSITE; PS00886; TRBD; 1.
KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW Protein biosynthesis; RNA-binding; tRNA-binding.
FT SITE 12 22 "HIGH" region.
FT SITE 308 312 "KMSKS" region.
FT DOMAIN 562 665 tRNA-binding.
FT BINDING 311 311 ATP (By similarity).
SQ SEQUENCE 665 AA; 74884 MW; AD5140C0C91BCFE6 CRC64;
Query Match 82.8%; Score 2880.5; DB 1; Length 665;
Best Local Similarity 82.3%; Pred. No. 2.9e-164;
Matches 547; Conservative 51; Mismatches 64; Indels 3; Gaps 2;
QY 3 EKRFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDFVYLTGLDHEGQKIQKAE 62
Db 2 KKFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGHEVYLTGLDHEGQKIQTKAK 61
QY 63 RAGITPQAVDGMVGVKELWQLDLSYDKFIRTTDDYHEKVAQVFERLLAQDDIYLGE 122
Db 62 RAGITPQYVDNMKADVKALWQLDLSYDKFIRTTDDYHEKVAQVFERLLAQDDIYLGE 121
QY 123 YSGWYSVDSDEFFTESQLEAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDRLV 182
Db 122 YSGWYSVDSDEFFTESQLEAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYADRLV 181

QY 541 EKEWNPDEVELKLNKIDBIKFEDEPKVEIRVAEVEKSKVSGDKLLQFRILDAGDGEDROI 600
DB 541 EKEWNPDEVALKEKQVITETFDVAEIRVAEVEKSKVSGDKLLRFRVDAGDGDROI 600
QY 601 LSGIAKYYPNEQELVGKKQIVANLKPRQMKKYVSGMILSAEHGDKLTLLTVDPAVFN 660
DB 601 LSGIAKYYPNEQELVGKKQIVANLKPRQMKKYVSGMILSAEHGDKLTLLTVDSSVFN 660
QY 661 GSVIG 665
DB 661 GSIIG 665

RESULT 5
SYM_STRPB STANDARD; PRT; 665 AA.
AC QBP298;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 03-JUL-2004 (Rel. 44, Last annotation update)
DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)
DE (MetRS)
GN Name=metG; Synonyms=metS; OrderedLocusNames=spyM18_0468;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=MGAS232 / Serotype M18;
RA PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.B., Sturdevant D.B., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
CC -!- FUNCTION: Is required not only for elongation of protein synthesis
CC but also for the initiation of all mRNA translation through
CC initiator tRNA(fMet) aminoacylation (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
CC diphosphate + L-methionyl-tRNA(Met).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC MetG subfamily 2B.
CC -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE009986; AL97197.1; ..
CC HSP; P23395; IABH.
CC HAMAP; MF 01228; fused; 1.
CC InterPro; IPR004495; MetG Cterm.
CC InterPro; IPR008224; MetRS dimerising.
CC InterPro; IPR002300; tRNA-synt_1a.
CC InterPro; IPR001412; tRNA-synt_1.
CC InterPro; IPR002304; tRNA-synt_met.
CC InterPro; IPR003080; tRNA-synt_1a_bind.
CC Pfam; PF00133; tRNA-synt_1.
CC Pfam; PF01588; tRNA_bind_1.
CC PIRSF; PIRSF001528; MetRS dimerising; 1.
CC PRINTS; PR01041; TRNASYNTHET.
CC TIGRFAMs; TIGR00398; metG; 1.

TIGRFAMs; TIGR00399; metG_C term; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
DR PROSITE; PS00886; TRSD; 1.
KW Aminoacyl-tRNA synthetases; ATP-binding; Complete proteome; Ligase;
KW Protein biosynthesis; RNA-binding; tRNA-binding.
FT SITE 12 22 "HIGH" region.
FT SITE 308 312 "KSKS" region.
FT DOMAIN 562 665 tRNA-binding.
FT BINDING 311 311 ATP (By similarity).
SQ SEQUENCE 665 AA; 74869 MW; 4451CF7CF2A2D8 CRC64;

Query Match 82.5%; Score 2868.5; DB 1; Length 665;
Best Local Similarity 81.7%; Pred. No. 1.5e-163;
Matches 543; Conservative 55; Mismatches 64; Indels 3; Gaps 2;

QY 3 EKNFYITTPYPSGKLHIGSAITTIACDVLYARKLMGYDVFYLTGLDEHGKIQCAE 62
DB .
QY 2 KAPFYITTPYPSGKLHIGSAITTIACDVLYARKLMGHEVFYLTGLDEHGKIQCAE 61
DB .
QY 63 EAGITFQAYVDGMVGVKELMQLDLSYDKFIRTTDDYHEKVAQVFFRLAQDDIYLGE 122
DB 62 EAGISPTQYVDNMVAKVQALMQLDLSYDTFIRTTDDYHEEVAVAFKLLAQDDIYLGE 121
QY 123 YSGWYSVSDERFFTESQLAEVFRDEAGNVGTGIIAPSGHEVWSEESYFLRLSKYQDRLV 182
DB 122 YSGWYSVSDERFFTESQLAEVFRDEAGNVGTGIIAPSGHEVWSEESYFLRLSKYQDRLV 181
QY 183 EFKKAHPFITPDGRNLNEMLRNFIEPGLDELAVSRITFTWGPVPSNPKHVVYVWIDALL 242
DB 182 AFFKERPDFIQPDGEMNEWKNFIEPGLDELAVSRITFTWGPVPSNPKHVVYVWIDALL 241
QY 243 NVATAGVAQDEHGNFDFXNGTVFHMVQKDIILRHISYIWPILLMMLDKVLPDRLIAHGW 302
DB 242 NVATAGYGOANHANFDFXNGTVFHMVQKDIILRHISYIWPILLMMLDKVLPDRLIAHGW 301
QY 303 FVWKDGKMSKSGNVVYVPEMLVERVGLDPLRYVLMRNLPVSGDGTFTTPEYVGRINVELA 362
DB 302 FVWKDGKMSKSGNVVYVPEMLVERVGLDPLRYVLMRNLPVSGDGTFTTPEYVGRINVELA 361
QY 363 NDGLNLLKRTYSMINKYFDGQIPAYVEGVTEFDHVLAEVAKSIADPHTHMEAVDYPRAL 422
DB 362 NDGLNLLKRTYAMINKYFDGQIPAYVNDNGTAFDADLSQLIDAQADYHKHMEAVDYPRAL 421
QY 423 EAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHWQASIRVVAHLISPEMETSRA 482
DB 422 EAVWTLIARTNKYIDETAPWVLAKEDGDKQALASVMAHLASLRVVAHVIQPFMMETSAA 481
QY 483 VLTQIGLEEVSSLENLSLADFPADVTVVAKGTPPPRLNMEEEIAYIKQMEGNKPAY-- 540
DB 482 IMAQGLAPVSDLSLALADFPANTKVAKGTPPPRLDMEAEIDYIKAQMGDSSAISQ 540
QY 541 EKEWNPDEVELKLNKIDBIKFEDEPKVEIRVAEVEKSKVSGDKLLQFRILDAGDGEDROI 600
DB 541 EKEWNPDEVALKEKQVITETFDVAEIRVAEVEKSKVSGDKLLRFRVDAGDGDROI 600
QY 601 LSGIAKYYPNEQELVGKKQIVANLKPRQMKKYVSGMILSAEHGDKLTLLTVDPAVFN 660
DB 601 LSGIAKYYPNEQELVGKKQIVANLKPRQMKKYVSGMILSAEHGDKLTLLTVDSSVFN 660
QY 661 GSVIG 665
DB 661 GSIIG 665

RESULT 6
QSDYDO PRELIMINARY; PRT; 665 AA.
ID QSDYDO
AC QSDYDO;
DT 01-NAR-2003 (TremBLrel. 23, Created)
DT 01-NAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-NAR-2004 (TremBLrel. 26, Last annotation update)
DE Methionyl-tRNA synthetase.
GN Name=metG; OrderedLocusNames=SAG1557;

OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

OX NCBI_TaxID=216466;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=2603 V/R / Serotype V;

RX MEDLINE=2222298; PubMed=12200547; DOI=10.1073/pnas.182380799;

RA Tettelin H., Wasiņani V., Cieleszczak M.J., Eisen J.A., Peterson S.N.,

RA Wessels M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,

RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,

RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,

RA Radune D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,

RA Carthy H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,

RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D.,

RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,

RA Fraser C.M.;

RT "Complete genome sequence and comparative genomic analysis of an

RT emerging human pathogen, serotype V Streptococcus agalactiae.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).

DR EMBL; AEO14263; AAN00423.1; -.

DR HSSP; P23395; 1A8H.

DR TIGR; SAG1557; -.

DR GO; GO:0005737; Cytoplasm; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0004825; F:methionine-tRNA ligase activity; IEA.

DR GO; GO:0000049; F:RNA binding; IEA.

DR GO; GO:0006431; P:methionyl-tRNA aminoacylation; IEA.

DR GO; GO:0006412; P:protein biosynthesis; IEA.

DR InterPro; IPR004495; MetG Cterm.

DR InterPro; IPR002300; tRNA-synt_la.

DR InterPro; IPR002304; tRNA-synt_met.

DR InterPro; IPR009080; tRNA-synt_la_bind.

DR InterPro; IPR002547; tRNA bind.

DR Pfam; PF00133; tRNA-synt_I; 1.

DR Pfam; PF01588; tRNA_bind; 1.

DR PRINTS; PR01041; TRNASYNTHET.

DR TIGRfams; TIGR00398; metG; 1.

DR TIGRfams; TIGR00399; metG Cterm; 1.

DR PROSITE; PS00886; TRBD; 1.

KW Aminoacyl-tRNA synthetase; Complete proteome.

KW SEQUENCE 665 AA; 75589 MW; 7C54FE064D053C1D CRC64;

Query Match 78.5%; Score 2732.5; DB 2; Length 665;

Best Local Similarity 78.2%; Pred. No. 2.2e-155;

Matches 519; Conservative 61; Mismatches 81; Indels 3; Gaps 3;

3 EKNNFYITPIYPSGKLHIGSAVTTIACDVLARYKELMGYDYVLTGLDHEGQKIQQXAE 62

4 KKSFYITPIYPSGKLHIGSAVTTIACDVLARYKELMGYDYVLTGLDHEGQKIQQXAE 63

63 EAGITPOAYVDGMAVGKELQWLLDISYKFKRTTDDYHEKVAQVFERLLAODDIYLGE 122

64 EAGITPOAYVDGMAVGKELQWLLDISYKFKRTTDDYHEKVAQVFERLLAODDIYLGE 123

123 YSGWYSVSDDEFFTESQLAEVFRDEAGNTGGTAPSGHEVWVSSESYFLRLSKYQDRILV 182

124 YSGWYSVSDDEFFTESQLAEVFRDEAGNTGGTAPSGHEVWVSSESYFLRLSKYQDRILV 183

183 EFPKAPHEFITDGRNEMLNRFIEPGLDLAVSRFTTGWVPSNPKHVVYVWIDALL 242

184 AYAEHPEFITDGRNEMLNRFIEPGLDLAVSRFTTGWVPSNPKHVVYVWIDALL 243

243 NYATALGYA-QDEHGDFNFGWCTVFMVKGDTLRFHSIYWPILLMLDVKLPDRLLAHG 301

244 NYATALGYA-QDEHGDFNFGWCTVFMVKGDTLRFHSIYWPILLMLDVKLPDRLLAHG 302

302 WFWKDGKMSKGNVYVPEMLVERVGLDPLRYLMLNLPVSGDGTFTPDYVGRINVEL 361

303 WFWKDGKMSKGNVYVPEMLVERVGLDPLRYLMLNLPVSGDGTFTPDYVGRINVEL 362

362 ANDLGNLLNRTIAMVVKYFDEGVPRFAV-ATDFADLASVATDSIENYKQMEAVDVPRA 421

Db 363 ANDLGNLLNRTIAMVVKYFDEGVPRFAV-ATDFADLASVATDSIENYKQMEAVDVPRA 421

Qy 422 LEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVWSHQASIRVVAHLIEPFFMMETSR 481

Db 422 LEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVWSHQASIRVVAHLIEPFFMMETSD 481

Qy 482 AVTQTGLEEVSSLENLSLADFPADVTVVAKGTPFPRLNMEETAIYIKQMEGNKPAVE 541

Db 482 AIMEQLGLGATFDEKLTADLPFEGVRVAKGSPFPRLDMEDEITYIKQMEGNKPAVE 541

Qy 542 KENNPDEVELKLNKDEKPEDPKVEIRVAEYKVEKSGDKLLQFLRDAGDGEDRQIL 601

Db 542 KENNPDEVELKLNKDEKPEDPKVEIRVAEYKVEKSGDKLLQFLRDAGDGEDRQIL 601

Qy 602 SGIAKYYPNEQELVGVKQIVANLKPRKMKYVSGMILSAEHDGKLTLLTVDPAPVNG 661

Db 602 SGIAKYYPNEQELVGVKQIVANLKPRKMKYVSGMILSAEHDGKLTLLTVDPAPVNG 661

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

Qy 662 SVTG 665

Db 662 SVTG 665

```

CC initiator tRNA(fMet) aminoacylation (By similarity).
CC -I- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
CC diphosphate + L-methionyl-tRNA(Met).
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC MetG subfamily 2B.
CC -I- SIMILARITY: Contains 1 tRNA-binding domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE006312; AA004890.1; -.
DR PIR; H86723; H86723.
DR HSP; P23395; 1A8H.
DR HAMAP; MF_01228; fused; 1.
DR InterPro; IPR004435; MetG_Cterm.
DR InterPro; IPR008224; MetRS_dimerising.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002304; tRNA-synt_met.
DR InterPro; IPR009080; tRNA_synth_1a_bind.
DR InterPro; IPR002547; tRNA_bind.
DR Pfam; PF00133; tRNA-synt_1; 1.
DR Pfam; PF00188; tRNA_bind; 1.
DR PIRSF; PIRSF001528; MetRS_dimerising; 1.
DR PRINTS; PR01041; TRNASYNTHET.
DR TIGRFAMs; TIGR00398; metG; 1.
DR TIGRFAMs; TIGR00399; metG_Cterm; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
DR PROSITE; PS00866; TRSD; 1.
KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW Protein biosynthesis; RNA-binding; tRNA-binding.
FT SITE 14 24 "HIGH" region.
FT SITE 308 312 "KMSKS" region.
FT DOMAIN 559 662 tRNA-binding.
FT BINDING 311 311 ATP (By similarity).
FT BINDING 662 AA; 75777 MW; 5536AAAD6F1CB91 CRC64;
SQ SEQUENCE 662 AA; 75777 MW; 5536AAAD6F1CB91 CRC64;
-----
Query Match 68.8%; Score 2392.5; DB 1; Length 662;
Best Local Similarity 68.6%; Pred. No. 5.3e-135;
Matches 458; Conservative 88; Mismatches 113; Indels 9; Gaps 7;
-----
Qy 1 MSE-KNFYITPIYPSPGKLHIGSAYTIACDVLARYKRLMGVDVFLTGLDHEGSKIQ 59
Db 1 MTENKTFYITPIYPSPGKLHLGSSYTIACDVLARYKRLMGVDFDTFLTGLDHEGKLIQR 60
-----
Qy 60 KAEAGITPOAYVDGMAVGKELWQLLDDISYDKFIRTTDDYHEKVAQVFERLLAQDDIY 119
Db 61 KAEELGWTPKXYLDPMADVQELWKLLDISYDKFIRTTDYHEAAVAKAFQLLEQDDIY 120
-----
Qy 120 LGEYSGHYVSDBEFFTESQALAVFRDEAGNVTCGIAPSGHEVWVSEESYFLRSKYQD 179
Db 121 LGYAGHYVSDBEFFTQLEEIFRDESGNITGGIAPSGHEVWVSEETFFFRMGKYAD 180
-----
Qy 180 RLVEFFKAKHPEFTIPDGRLEMLRNFTPEGLDLAVSRITFTWGVPPSPNEKHVVYWMID 239
Db 181 WLQOYDEHDPFTQPEVRKEMWNNFKPGLDGLALTFTSFTWGIPIVPSNPKHVYVWF 240
-----
Qy 240 ALLNYATALGYAOEHOHGFDPKFNWGVDFVHMVGKDILRFHSIYWIPILLMWLDVKVLPDRLLIA 299
Db 241 ALLNYITALGVNSDNDNFKKYWPNG--INNVGKEIVRFHTIYIYIIMLHALGLPAPKKIFA 298
-----
Qy 300 HGVFVMDKGWKSXSGKNVYVPEMLVARYGLDPLRYIYLRNLVPGSDGFTTDPEDYVGRINY 359
Db 299 HGWLMLKDGWKSXSGKNVYVPEMLIERYGLDAVRIYLRNLVPGSDGFTTDPEDYVGRINFP 358

```


Db 600 NHRQILSGIAEFYDPDPAALGKKVIVANLKPRK-MRGQISQGMILSAESPEGLQIWEA 658

QY 655 DPAVPNGSVI 664

Db 659 PKEMPNGAGI 668

RESULT 10

SYM_OCEIH STANDARD; PRT; 673 AA.

AC P59079;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase) (MetRS).

GN Name=metG; OrderedLocusNames=OB0046;

OS Oceanobacillus ihewensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.

OK NCBI_TaxID=182710;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H7E831 / DSM 14371 / JCM 11309;

RX MEDLINE=2220767; PubMed=12235376;

PA Takami H., Takaki Y., Uchiyama I.;

RT "Genome sequence of Oceanobacillus ihewensis isolated from the Iheya Ridge and its unexplored adaptive capabilities to extreme environments."

RL Nucleic Acids Res. 30:3927-3935(2002).

CC -!- FUNCTION: Is required not only for elongation of protein synthesis but also for the initiation of all mRNA translation through initiator tRNA(fMet) aminoacylation (By similarity).

CC -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP + diphosphate + L-methionyl-tRNA(Met).

CC -!- SUBUNIT: Homodimer (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family. MetG subfamily 2B.

CC -!- SIMILARITY: Contains 1 tRNA-binding domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AP004593; BAC12002.1; -.

CC HSSP; P23395; 1A8H.

CC HAMAP; MF_01228; fused; 1.

CC InterPro; IPR004495; MetG_Cterm.

CC InterPro; IPR008224; MetRS_dimerising.

CC InterPro; IPR008994; Nucleic_acid_OB.

CC InterPro; IPR002300; tRNA-synt_1a.

CC InterPro; IPR001412; tRNA-synt_1.

CC InterPro; IPR002304; tRNA-synt_met.

CC InterPro; IPR003080; tRNA-synt_la_bind.

CC InterPro; IPR002547; tRNA_bind.

CC Pfam; PF00133; tRNA-synt_1; 1.

CC Pfam; PF01588; tRNA_bind; 1.

CC PIRSF; PIRSF001528; MetRS_dimerising; 1.

CC PRINTS; PR01041; TRNASYNTMET.

CC TIGRFAMs; TIGR00398; metG; 1.

CC TIGRFAMs; TIGR00399; metG_Cterm; 1.

CC PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.

CC PROSITE; PS00886; TRSD; 1.

KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase; Protein biosynthesis; RNA-binding; tRNA-binding.

FT SITE 14 24 "HIGH" region.

FT SITE 310 314 "KMSKS" region.

FT DOMAIN 571 673 tRNA-binding.

FT BINDING 313 313 ATP (By similarity).

SQ SEQUENCE 673 AA; 76610 MW; D6E8EF706BD9BCAD CRC64;

Query Match 64.1%; Score 2230.5; DB 1; Length 673;

Best Local Similarity 62.6%; Pred No. 2.8e-125;

Matches 420; Conservative 100; Mismatches 140; Indels 11; Gaps 6;

QY 3 EKNFYITPIYYPGSKLHIGSAITTIACDLARYRLMGVDFYLTGDEHGQKIQOAK 62

Db 4 ENAFYITPIYYPGSKLHIGNAITTIACDVMARYRMGRGDFYLTGDEHGQKIQOAK 63

QY 63 EAGITPOAVDGMVGVKELMQLLDISYDKFTRTDDYHEKVVAOVFERLLAQQDIYLGE 122

Db 64 ENNISFKAYVDDMAEGMKLWNTLEISNDKFIITTEHQKKVVADIFERFEGQDIYDE 123

QY 123 YSGWYSVDEEFTTSQALAEVERDEAGNVGTGIAP-SGHEVEMVSESEYFRLSKYQRL 181

Db 124 YEGWYSVPDETETETQLEDVERDDGNVIGKSPDGHVPVELIKEESYFFRMSKYADRL 183

QY 182 VEPFKAHEFITPDGLNAEMLENFIERGLEDAVSTTFTWGPVPSPKVVVYVWIDAL 241

Db 184 LKPYEDNPEFIOPESRKNNMNNFKPGLLEDVAVSRITTFSGVQVPSKPKVVVYVWIDAL 243

QY 242 LNVATALGYAODEHGNFDFKFWNGTTFVHMVGDKILFHSIYWPILMLMLDVKLPDLRIAG 301

Db 244 THYITALGYGSEDTSLYDKFWPADV-HMVKEIVAFHTIYWPILMALDLPLPKKVFAG 302

QY 302 WFMKDGKMSKSGNVVYPEMLVERYGLDPLRYIMRMLPVGSDGTFTFPEDYVGRINTEL 361

Db 303 WLLMKDGKMSKSGNVVYPEMLVERYGLDALRYILMRVAFVGSFGVFTPEDFISRVNYDL 362

QY 362 ANDLGNLNRVTSMYKVFYDGOIPAVYGVTFDFHVLAEVAKSIADFHHEAVDYDPA 421

Db 363 ANDLGNLNRVTAMINKYFDGKVPFKGEVTSFDGELQTTANNAVKEKHEGQVFSDA 422

QY 422 LEAVMTLSRTNKYIDETAPWVLDKDEALRDQASVMSHQASIRVVAHLIEPFFMETS 481

Db 423 LKQWILLISRANKYIDETEPWIVAKDEGRNELASVWHLAESLHAAALMLQFLTHAPK 482

QY 482 AVLTOIGLEEVSSLE--NLSLADFPADTVVAKGTFPIPRLNWEIEIYKQMEGNKPA 539

Db 483 KIAQGLGIEEYGLDWTGTFGNFENTTVKKGTFPIPRLLDDEAAIYRDMANGANA 542

QY 540 VEKE-----WNPDEVELKINKD-EIKFEDFDKVEIRVAEVEKSVKVGSDKLLQFRLDAG 593

Db 543 ASSEDETGDWDFNETDLVSEKEKQIKYDFDKVVKVAEKDCSVKVGADKLLKFLDAG 602

QY 594 DEEDQILSGIAKYIPNEQELVKKYQIVANLKPRKMKKYVQSGMILSAEHDGKLTLLT 653

Db 603 DNGDRQILSGIAEYYSPEQLIGKKVIVANLKPRK-MRGEISQGMILSAEYDGLQIWE 661

QY 654 VDPVAVPNGSVI 664

Db 662 APSEAPNGSSI 672

RESULT 11

SYM_LACPL STANDARD; PRT; 683 AA.

AC Q88Z97;

DT 29-MAR-2004 (Rel. 43, Created)

DT 29-MAR-2004 (Rel. 43, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase) (MetRS).

GN Name=metG; Synonym=metS; OrderedLocusNames=lp_0454;

OS Lactobacillus plantarum.

OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;

OC Lactobacillus.

OK NCBI_TaxID=1590;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NCIMB 8826 / WCPSL;

RX MEDLINE=22480296; PubMed=12566566; DOI=10.1073/pnas.0337704100;

RA Kleebebe M., Boekhorst J., van Kranenburg R., Molenaar D.,
 RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
 RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
 RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
 RA De Vos W.M., Siezen R.J.,
 RT "Complete genome sequence of *Lactobacillus plantarum* WCFS1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
 CC -i- FUNCTION: Is required not only for elongation of protein synthesis
 CC but also for the initiation of all mRNA translation through
 CC initiator tRNA(fMet) aminoacylation (By similarity). AMP +
 CC -i- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
 CC diphosphate + L-methionyl-tRNA(Met).
 CC -i- SUBUNIT: Homodimer (By similarity).
 CC -i- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -i- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
 CC MetG subfamily 2B.
 CC -i- SIMILARITY: Contains 1 tRNA-binding domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL935253; CAD63085.1; -;
 CC HSPSP; Q9V011; IMKH.
 DR HAMAP; MF_01228; -; 1.
 DR InterPro; IPR004495; MetG Cterm.
 DR InterPro; IPR008994; Nucleic_acid_OB.
 DR InterPro; IPR002300; tRNA-synt_1a.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002304; tRNA-synt_met.
 DR InterPro; IPR009080; tRNA-syn_1a_bind.
 DR InterPro; IPR002547; tRNA_bind.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR Pfam; PF01588; tRNA_bind; 1.
 DR PRINTS; PR01041; TRNASYNTHET.
 DR PROSITE; PS00178; AA TRNA LIGASE I; 1.
 DR PROSITE; PS50886; TRBD; 1.
 KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
 KW Protein biosynthesis; tRNA-binding.
 FT SITE 15 25 "HIGH" region.
 FT SITE 311 315 "KMSKS" region.
 FT DOMAIN 581 683 tRNA-binding.
 FT BINDING 314 314 ATP (By similarity).
 SQ SEQUENCE 683 AA; 76796 MW; 2E68E67D28DA8560 CRC64;
 Query Match 63.3%; Score 2201; DB 1; Length 683;
 Best Local Similarity 61.4%; Pred. No. 1.7e-123;
 Matches 416; Conservative 106; Mismatches 137; Indels 20; Gaps 7;
 QY 2 SERNFYITPTIYPYSGKLHGSAITTIACDVLARYKRLMGDYVYLTGDERGKIQKQA 61
 DB 4 TKPTVITPTIYPYSGKLHGSAITTIACDVLARYKRLMGDYVYLTGDERGKIEEKA 63
 QY 62 BEAGIPQAVDGMVGVKELWLLDISYDKFTTTDDYHKVAVQVFERLLAQDDIYLG 121
 DB 64 EKLTNTPKSYVDGMVAKIQDLWLLSINDKFRITTTDDYHRAVQVIFRLLKNGDIYUG 123
 QY 122 EYSGWTSVSDEEFTBTSQLAVERFDAGNVGTGAGIAPSHEVWVSDESYFLRLSKYQDRL 181
 DB 124 EYSGWTSVDDEEYFTETQLAEVFRDDNGKVIQKAPSGHEVELVKEQSYFFKMSKYADWL 183
 QY 182 VEFKAPHEITPDGRNLNEMLRNFIPEGLDLAVSTTTTGWVPVSPNKHVYVWIDAL 241
 DB 184 LDYQSHPDIFIEPANRTEMINNFIPKGLDLAVSTSTGWVPVNSDPKHVYVWIDAL 243
 QY 242 LNVATALGYAQ-DEHGNFDFNMGTGVFHVGVGDKILRFHSTIYWPILMLMDVKLPORLIAH 300
 DB 244 TNYITALGYATGSEDLFNKFWADV-QMVGKEIVRFHYIYWPILHALGLPLPKKVFCH 302

QY 301 GWFVKDGRKMSKGNVYVPEMLVERYGLDPLRYILMRNLPGVSDGTFTPTEDYVGRINYE 360
 DB 303 GLLMKDGRKMSKGNVYVPEMLVERYGLDPLRYILMRNLPGVSDGTFTPTEDYVGRINYE 362
 QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTFEDHVLAEVAEKSADFEHTEANVDYPR 420
 DB 363 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTFEDHVLAEVAEKSADFEHTEANVDYPR 422
 QY 421 ALAEVMTLISRTNKYIDETAPWVLDDKDEALR--DQLASVMHMQASIRVVVAHLIEPFME 478
 DB 423 ALAEVMTLISRTNKYIDETAPWVLDDKDEALR--DQLASVMHMQASIRVVVAHLIEPFME 482
 QY 479 TSRAVLFOIGLEEVSLLENLSADPPADVTVVAKGTPPPRPLNMBEETAYIKQY---- 533
 DB 483 APKEIFTQLGLDPATLAIADLQADLPAGAQQVAVKGTPIPPRVDMDAEVEFLKGRMKTSD 542
 QY 534 --EGNK-----PAVEKEMNPDEVELKLNKDEIKFEDPKVEIRVAEVEKYSKVEGSD 583
 DB 543 KQGRKAMENAKHEAEVEQGNPAETNLTNLTPTAITIDDFDKVELKVAEVIITVKLKGAD 602
 QY 584 KLLQFRLDAGDGEDRQILSGIAYIPNEQELVGKQVQIVANLKPROMKKYKQGMILSA 643
 DB 603 KLLQFRLDAGDGEDRQILSGIAYIPNEQELVGKQVQIVANLKPROMKKYKQGMILSA 661
 QY 644 BHDGKLTLLTVDPVAVNGSVI 564
 DB 662 BHDGQVQLITVPDMMVNGSLI 682
 RESULT 12
 Q724N6 PRELIMINARY; PRT; 665 AA.
 ID Q724N6
 AC Q724N6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Methionyl-tRNA synthetase (EC 6.1.1.10).
 GN Name=metG; OrderedLocustNames=LMOF2365.0188;
 OS Listeria monocytogenes (serotype 4b / strain F2365).
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 CX NCBI_TaxID=265669;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15115801; DOI=10.1093/nar/gkh562;
 RA Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,
 RA Kolonay J.F., Rasko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,
 RA Peterson J.D., White O., Nelson W.C., Nierman W.C., Beanan M.J.,
 RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
 RA Haft D.H., Selengut J., Van Aken S.B., Khouri H.M., Fedorova N.,
 RA Forberger H.A., Tran B., Katharicu S., Wonderling L.D., Ulrich G.A.,
 RA Bayles D.O., Luchansky J.B., Fraser C.M.,
 RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
 RT borne pathogen *Listeria monocytogenes* reveal new insights into the
 RT core genome components of this species.";
 RL Nucleic Acids Res. 32:2386-2395(2004).
 DR EMBL; AE017322; AAT02975.1;
 DR GO; GO:0004825; F:ligase activity; IEA.
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR InterPro; IPR004495; MetG Cterm.
 DR InterPro; IPR008224; MetRS_dimerising.
 DR InterPro; IPR008994; Nucleic_acid_OB.
 DR InterPro; IPR002300; tRNA-synt_1a.
 DR InterPro; IPR002304; tRNA-synt_met.
 DR InterPro; IPR009080; tRNA-syn_1a_bind.
 DR InterPro; IPR002547; tRNA_bind.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR Pfam; PF01588; tRNA_bind; 1.
 DR PIRSP; PIRSF001528; MetRS_dimerising; 1.
 DR PRINTS; PR01041; TRNASYNTHET.
 DR TIGRFAMs; TIGR00398; metG; 1.
 DR TIGRFAMs; TIGR00399; metG_Cterm; 1.
 DR PROSITE; PS50886; TRBD; 1.
 KW Aminoacyl-tRNA synthetase; Complete proteome; Ligase.

```
SO SEQUENCE 665 AA; 75767 MW; B4FCE15778689C27 CRC64;
Query Match 60.8%; Score 2114; DB 2; Length 665;
Best Local Similarity 60.9%; Pred. No. 2,7e-118;
Matches 408; Conservative 98; Mismatches 146; Indels 18; Gaps 8;
Qy 3 EKN-FYITPIYPSGKHLHIGSAVTTIACDVLARYKRLMGYDVFYLTGLDEHGQKIQA 61
Db 5 EKNFYITPIYPSGKAHIGHAYTTVAGDAMARYKRLGYDVFYLTGLDEHGQKIQA 64
Qy 62 EAGITPQAYVDMGAVGKELWQLLDISYDKPIRTTDDYHEKVAQVFERLLAQDDIYLG 121
Db 65 KERGISEQEVYDEIAEGFQELWKLKLEISNTDFIRTTQDRHKTSVEKIFEQLLEQGIYLG 124
Qy 122 EYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSSEYFLRLSKYQDRL 191
Db 125 EYEGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSSEYFLRLSKYQDRL 194
Qy 182 VEFKKAPEITPDGRNLNEMRNFIPEGLDLAVSRITFTWGVPSNPKHVVYVMDAL 241
Db 185 VEYNSHPFELPESRKNEINNFIKPGLEDLAVSRITFTWGVPSNPKHVVYVMDAL 244
Qy 242 LNYATALGYAQDEHGNFDFKNGTVFHMVGKDILRFHSIYWPILLMMLDVKLPDRLIAHG 301
Db 245 SNY-TALGYNTDNTKFKYWPADV-QIVGKEIVRFHTIYWPIMLMALDLPKQVFGHG 303
Qy 302 WFWKDGKMSKGNVYVPEMLVERYGLDPLRYILMENLPVGS DGTFTPDYGRINVEL 361
Db 304 WILMKDGKMSKGNVYVPEMLIDRYGLDALRYLLREVFGSDGLFTPDYGRINVEL 363
Qy 362 ANDLGNLLNRTVAMINKYFDGEIPAYQGNVTEFDQTLVDPKNNVKEYESGDMHQFSA 421
Db 364 ANDLGNLLNRTVAMINKYFDGEIPAYQGNVTEFDQTLVDPKNNVKEYESGDMHQFSA 423
Qy 422 LEAVNTLSRTNKYIDETAPWLDKDEALRDQASVMSHQASIRVVAHLIEPFMMETSR 481
Db 424 LNQLWSLISRTNKYIDETAPWLDKDEALRDQASVMSHQASIRVVAHLIEPFMMETSR 483
Qy 482 AVLTOIGLEEVSSLENL-----SLADFPADVTVVAKGTPIPRLNMEEEIAYIKEOME 534
Db 484 EIFLQGLQE-----ENLKKWDSIYGYGEIPAGTIVVKKGTPIPRLEAEVEVYIQDEM 539
Qy 535 GNKPAVEKWNPDVELKLNKDEIKFEDPKVEIRVAEVEKVSKEGSDKLLQFRLDAGD 594
Db 540 GSAPAPAKE--AAEVE-ALETPOIGIEDFDKIDLRVAEVEKQVKKADKLLCFQLDLGE 596
Qy 595 GEDROILSGIAKYYPNEQELVGKQIVANLKPKRMKKYVSGMILSAEHGDKLTLLTV 654
Db 597 GKLRQVLSGIAEYFQPE-ELIGKKVIVVSNLKPVK-LRGLMSEGMILSGEKGKLSVIEA 654
Qy 655 DPAPNGSVI 664
Db 655 SSALPNGAKV 664
RESULT 13
AA020975 PRELIMINARY; PRT; 665 AA.
AC AA020975;
DT 10-MAY-2004 (Tremblrel. 27, Created)
DT 10-MAY-2004 (Tremblrel. 27, Last sequence update)
DB Methionyl-tRNA synthetase. 27, Last annotation update)
GN METG OR LMOF2365.0188.
OC Listeria monocytogenes str. 4b F2365.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria;
OC Listeria monocytogenes.
OX NCBI_taxid=265669;
[1]
FN SEQUENCE FROM N.A.
RP STRAIN=4b F2365;
RC PubMed=15115801;
RA Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,
```

```
RA Kolonay J.F., Rasko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,
RA Peterson J., White O., Nelson W.C., Nierman W.C., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
RA Haft D.H., Selengut J., Van Aken S., Khouri H., Fedorova N.,
RA Forberger H., Tran B., Kathariou S., Wonderling L.D., Uhlir G.A.,
RA Bayles D.O., Luchansky J.B., Fraser C.M.;
RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
RT borne pathogen Listeria monocytogenes reveal new insights into the
RT core genome components of this species.";
RL Nucleic Acids Res. 32:2386-2395(2004).
DR EMBL; AB017322; AA020975.1; -.
KW Aminoacyl-tRNA synthetase; Ligase.
SQ SEQUENCE 665 AA; 75767 MW; B4FCE15778689C27 CRC64;
```

```
Query Match 60.8%; Score 2114; DB 2; Length 665;
Best Local Similarity 60.9%; Pred. No. 2,7e-118;
Matches 408; Conservative 98; Mismatches 146; Indels 18; Gaps 8;
Qy 3 EKN-FYITPIYPSGKHLHIGSAVTTIACDVLARYKRLMGYDVFYLTGLDEHGQKIQA 61
Db 5 EKNFYITPIYPSGKAHIGHAYTTVAGDAMARYKRLGYDVFYLTGLDEHGQKIQA 64
Qy 62 EAGITPQAYVDMGAVGKELWQLLDISYDKPIRTTDDYHEKVAQVFERLLAQDDIYLG 121
Db 65 KERGISEQEVYDEIAEGFQELWKLKLEISNTDFIRTTQDRHKTSVEKIFEQLLEQGIYLG 124
Qy 122 EYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSSEYFLRLSKYQDRL 181
Db 125 EYEGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSSEYFLRLSKYQDRL 184
Qy 182 VEFKKAPEITPDGRNLNEMRNFIPEGLDLAVSRITFTWGVPSNPKHVVYVMDAL 241
Db 185 VEYNSHPFELPESRKNEINNFIKPGLEDLAVSRITFTWGVPSNPKHVVYVMDAL 244
Qy 242 LNYATALGYAQDEHGNFDFKNGTVFHMVGKDILRFHSIYWPILLMMLDVKLPDRLIAHG 301
Db 245 SNY-TALGYNTDNTKFKYWPADV-QIVGKEIVRFHTIYWPIMLMALDLPKQVFGHG 303
Qy 302 WFWKDGKMSKGNVYVPEMLVERYGLDPLRYILMENLPVGS DGTFTPDYGRINVEL 361
Db 304 WILMKDGKMSKGNVYVPEMLIDRYGLDALRYLLREVFGSDGLFTPDYGRINVEL 363
Qy 362 ANDLGNLLNRTVAMINKYFDGEIPAYQGNVTEFDQTLVDPKNNVKEYESGDMHQFSA 421
Db 364 ANDLGNLLNRTVAMINKYFDGEIPAYQGNVTEFDQTLVDPKNNVKEYESGDMHQFSA 423
Qy 422 LEAVNTLSRTNKYIDETAPWLDKDEALRDQASVMSHQASIRVVAHLIEPFMMETSR 481
Db 424 LNQLWSLISRTNKYIDETAPWLDKDEALRDQASVMSHQASIRVVAHLIEPFMMETSR 483
Qy 482 AVLTOIGLEEVSSLENL-----SLADFPADVTVVAKGTPIPRLNMEEEIAYIKEOME 534
Db 484 EIFLQGLQE-----ENLKKWDSIYGYGEIPAGTIVVKKGTPIPRLEAEVEVYIQDEM 539
Qy 535 GNKPAVEKWNPDVELKLNKDEIKFEDPKVEIRVAEVEKVSKEGSDKLLQFRLDAGD 594
Db 540 GSAPAPAKE--AAEVE-ALETPOIGIEDFDKIDLRVAEVEKQVKKADKLLCFQLDLGE 596
Qy 595 GEDROILSGIAKYYPNEQELVGKQIVANLKPKRMKKYVSGMILSAEHGDKLTLLTV 654
Db 597 GKLRQVLSGIAEYFQPE-ELIGKKVIVVSNLKPVK-LRGLMSEGMILSGEKGKLSVIEA 654
Qy 655 DPAPNGSVI 664
Db 655 SSALPNGAKV 664
```

```
RESULT 14
SYM_LISMO STANDARD; PRT; 664 AA.
AC Q8YAF2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
```

	Matches	407; Conservative	101; Mismatches	144; Indels	18; Gaps
QY	3	EKN-FYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDYFYITLGDHGQKIQQA	61		
Db	4	EKNFYITPIYPSGKAHIGHAVTTVAGDAMARYKRLKGYDVFYITLGDHGQKIQQA	63		
QY	62	EEAGTTPQAVDGVAGVYKELWQLLDISYDKFIRTTDDVHEKVVAAQVBERLLAODDIYLG	121		
Db	64	KERLSGEQYVDEIABGFQELWKKLEISNTDFIRTTQDRHKTSEVKIFQGLEQGDIIYG	123		
QY	122	EYSGWYSVDSDEFFTESQLAEVRDEAGNVTGGIAPSGCHEVWVSSESYFLRLSKYODRL	181		
Db	124	EYEGWYSVDSDEYFTETQLEBEVYKDENGKVIIGKAPSGNEVELVKESYFFRMSKYADRL	183		
QY	192	VEFFKAMPERTTPGRLNEMLRNFIERGELEDIAVSRTTFTWGVVPSPNKHVVYVWIDAL	241		
Db	184	VEYNSHPFEFLTPSRKNEMINNFIKPLGSLDLAVSRTTFDWMGILKVPNGPHVVYVWIDAL	243		
QY	242	LNATALCYAQADEHNGFDKFWNGTVFHMVGKDILRFHSIYWPILLMMLDVKLPDLRIAHG	301		
Db	244	SNYITALGYNTDNTKQKYAPADV-QIVGKEIVRFHTIYWPIMLXALDLPLPKMVEGHG	302		
QY	302	WFVKDGMKSXSGNVVYPEWLVERYGDLPIRYILMRNLVPGSDGCTFTPEDYVGRINYL	361		
Db	303	WILKMGKMSXSGNVVDPYMLIDRYGLDARYLLREVPFGSDGLFTPEDFVDRYNFDL	362		
QY	362	ANDLGNLLNRVTSMINKYFDGQIPAYVEGTEFDHVLAEVAEKSIADFTHTMEADVYRA	421		
Db	363	ANDLGNLLNRIVAMINKYFDGEIPAYOGNTEFDQLIVDFKNNVKEYSGSDHMQFSVA	422		
QY	422	LEAVWTLISRTNKYIDTAPWLDKDBALBDQLASVNSHMQASIRVVAHLIIEPFMMETR	481		
Db	423	LNQLWSLISRTNKYIDETAPWALAKDBDKETELASVNTHLAENLRIIAVLLOPFLTRTPG	482		
QY	482	AVLTQIGLEEVSSLENL-----SLADFPADVTVAKGTPIPFPRLMNBEETAYIKEOME	534		
Db	483	EIFQLGLQE-----ENLKKWDSIYGEIEIPAGTIVVKGPPIEPRIDAEVEVTIODEMK	538		
QY	535	GKPKAVEKEMNPDEVELEKLANKDEIKFEDFDKVBIRVAEVAEKVSKVEGSKDLQFLRDAGD	594		
Db	539	GSAPAPAE--VAEVE-ALETPOIGIEDFDKIDLRVAEVKQVDKVKAKDKLCLCFQLDLGE	595		
QY	595	GEDQQLISGIAKYYPNQELVGKKVQIVANLKRPMKKVVSOGMTLSAEHDCKLTLVV	654		
Db	596	GKLRQVLSGIAEYFQPE-ELIGKKVIVVSNLKPVK-LRGLMSGMLSGEKDGKLSVIEA	653		
QY	655	DPAPVNGSVI	664		
Db	654	SSALPNGAKV	663		
		[1]			
RESULT 15					
SYM_LISIN	ID	SYM_LISIN	STANDARD;	PRT;	664 AA.
AC	Q92F90;				
DT	28-FEB-2003	(Rel. 41, Created)			
DT	28-FEB-2003	(Rel. 41, Last sequence update)			
DE	05-JUL-2004	(Rel. 44, Last annotation update)			
DE	Methionyl-tRNA synthetase (EC 6.1.1.10)	(Methionine--tRNA ligase)			
DE	(MetRS);				
DN	Name=metG; Synonym=metS; OrderedLocusNames=lin0216;				
OS	Listeria innocua				
OC	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.				
OX	NCBI_taxid=1642;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CLIP 11262 / Serovar 6a;				
RC	MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;				
RX	Glaser P., Franeul L., Buchrieser C., Rusniok C., Amend A.,				
RA	Baquerio F., Berche P., Bloecher H., Brandt P., Chakraborty T.,				
RA	Charbit A., Cretouani F., Couve E., de Daruvar A., Dehoux P.,				
RA	Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Duessegert O.,				
RA	Entian K.-D., Fishi H., Garcia-del Portillo F., Garrido P.,				

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voes H., Wehland J., Cossart P.,
RT "Comparative genomics of *Listeria species*.";
RL Science 294:849-852(2001).
CC -i- FUNCTION: Is required not only for elongation of protein synthesis
CC but also for the initiation of all mRNA translation through
CC initiator tRNA(fMet) aminoacylation (By similarity). = AMP +
CC -i- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
CC diphosphate + L-methionyl-tRNA(Met).
CC -i- SUBUNIT: Homodimer (By similarity).
CC -i- SUBCELLULAR LOCATION: Cytoplasmic.
CC -i- SIMILARITY: Belongs to Class-I aminoacyl-tRNA synthetase family.
CC MetG subfamily 2B.
CC -i- SIMILARITY: Contains 1 tRNA-binding domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL596163; CAC95449.1; .
CC PIR: A11459; A11459.
CC HSP: P21888; 1117.
CC Listlist; LIN00216; .
CC HAMAP; MF_01228; fused; 1.
CC InterPro; IPR004495; MetG_Cterm.
CC InterPro; IPR008224; MetRS dimerising.
CC InterPro; IPR002300; tRNA-synt_1a.
CC InterPro; IPR001412; tRNA-synt_1.
CC InterPro; IPR002304; tRNA-synt_met.
CC InterPro; IPR003080; tRNA-synt_1a_bind.
CC InterPro; IPR002547; tRNA_bind.
CC Pfam; PF00133; tRNA-synt_1; 1.
CC Pfam; PF01588; tRNA_bind; 1.
CC PIRSF; PIRSF001528; MetRS dimerising; 1.
CC PRINTS; PR01041; TRNASYNTHMET.
CC TIGRFAMs; TIGR00398; metG; 1.
CC TIGRFAMs; TIGR00399; metG_Cterm; 1.
CC PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
CC PROSITE; PS00886; TRBD; 1.
CC Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
CC Protein biosynthesis; RNA-binding; tRNA-binding.
CC SITE 15 25 "HIGH" region.
CC SITE 310 314 "KXSKS" region.
CC DOMAIN 563 664 tRNA-binding.
CC BINDING 313 313 ATP (By similarity).
CC SEQUENCE 664 AA; 75756 MW; 64A526B817B8195 CRC64;

Query Match 60.6%; Score 2107; DB 1; Length 664;
Best Local Similarity 60.6%; Pred. No. 7e-118;
Matches 406; Conservative 101; Mismatches 145; Indels 18; Gaps 8;

QY 3 EKN-FYITPTIYPSGKLHIGSAVTTIACDLVARYKLMGYDVPYLTGLDEHGQIOOKA 61
DB 4 EKNTFYITPTIYPSGKAHIGHAYTTVAGDAMARYKRLKGYDVPYLTGTDHGQIOAKA 63

QY 62 BEAGITPOAYVDGMAVGKELWQLLDISYDKFIETDDYHEKVAQVFERLLAQDDIYLG 121
DB 64 KERGISQEYVDETAEGFCELWKKLEISNTFIETQDRKTSVEKIFEQLEGGDIYLG 123

QY 122 EYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDRL 181
DB 124 EYEGWYSVDSDEYFTETQLEEVYKDENGKVIKKAPSGNEVELKESYFFRNSKYADRL 183

QY 182 VEFKAPHEFTTPDGRNLNMLNFIKPLEDLAVSRFTTFWGVPSNPKHVVYVWIDAL 241
DB 184 VEYNSHPFEFTLPBSRKNEMINNIKPLEDLAVSRFTTFWGIKVPGNPKHVVYVWIDAL 243
QY 242 LNYATALGYAODEHGNFDFKWNVTGFHVMVGKDIILRFHSIYWPILIMLMDVKLPORLIARH 301
DB 244 SNYITALGYNTDNDTKFKQYWPADV-QIVGKEIVRFHTIYWPIMLMALDLPKXWFGHG 302
QY 302 WFMKDGKXKSKSGNVVPEMLVRYGLDPLRYILMNLPLVSGSGTFTPEDYVGRINVEL 361
DB 303 WILMKDGKXKSKSGNVVDPYMLIDRYGLDALRYLLREVPFGSGLFTPEDFVDRVNDL 362
QY 362 ANDLGNLNRVTSMINKYFDGQIIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDYPRA 421
DB 363 ANDLGNLNRVTSMINKYFDGQIIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDYPRA 422
QY 422 LEAVWTLISRTNKYIDETAPVLDKDEALRDQLASVMSHWQASIRVVAHLIEPFWMTSR 481
DB 423 LNLWLSLISRTNKYIDETAPWALAKEEDKRTELASVMTHLAENLRIIAVLLQPLTRTPG 482
QY 482 AVLTOIGLEEVSSLENL-----SLADFPADVTVWAGTPIFFPRLNNEERIAYIKQME 534
DB 483 EIFLQGLQGE-----ENLKKWDSIVGYGPIPGTIVVKKGTPIFFALDRAEVEYTIQDEM 538
QY 535 GNKPAVEKEWNPDEVELKINKDEIKFEDFKVEIRVAEVEKSVKVGSDKLLQFLDAGD 594
DB 539 GSAPAPAE--TAEVE-ALETPQIGIEDFDKIDLRVAEVKQVDKVKADKLLCFOLDLGE 595
QY 595 GEDROILSGIAKYYPNEQELVKKQVIVANLKPROMKYYVSOGMILSAEHDGKLTILTV 654
DB 596 GKLRQVLSGIAEFYQPE-ELIGKKYIVVSNLKPVK-LRGLMSEGMILSGEXDKGKLSVIEA 653
QY 655 DPAVNGSVI 664
DB 654 NSALPNGAKV 663

Search completed: December 17, 2004, 21:59:30
Job time : 204 secs

This Page Blank (uspto)